Graduate Texts in Physics

Kurt Binder · Dieter W. Heermann

Monte Carlo Simulation in Statistical Physics

An Introduction

Sixth Edition

Springer

Graduate Texts in Physics

Graduate Texts in Physics

Graduate Texts in Physics publishes core learning/teaching material for graduate- and advanced-level undergraduate courses on topics of current and emerging fields within physics, both pure and applied. These textbooks serve students at the MS- or PhD-level and their instructors as comprehensive sources of principles, definitions, derivations, experiments and applications (as relevant) for their mastery and teaching, respectively. International in scope and relevance, the textbooks correspond to course syllabi sufficiently to serve as required reading. Their didactic style, comprehensiveness and coverage of fundamental material also make them suitable as introductions or references for scientists entering, or requiring timely knowledge of, a research field.

More information about this series at http://www.springer.com/series/8431

Kurt Binder · Dieter W. Heermann

Monte Carlo Simulation in Statistical Physics

An Introduction

Sixth Edition

Springer

Kurt Binder
Institut für Physik
Johannes Gutenberg Universität
Mainz, Germany

Dieter W. Heermann
Fakultät für Physik und Astronomie
Institut für Theoretische Physik
Ruprecht-Karls-Universität Heidelberg
Heidelberg, Germany

ISSN 1868-4513 ISSN 1868-4521 (electronic)
Graduate Texts in Physics
ISBN 978-3-030-10757-4 ISBN 978-3-030-10758-1 (eBook)
https://doi.org/10.1007/978-3-030-10758-1

Library of Congress Control Number: 2018968087

Preface to the Sixth Edition

The material covered in this book grew out of a series of lectures and complementary exercises presented at a Summer School held at Figueira da Foz (Portugal) in 1987. At that time, computer simulation methods started to find widespread applications; although one basic algorithm for Monte Carlo simulation in statistical physics, due to Metropolis et al., dates back to 1953, it found little use during the first 20 years after its invention. The reason for this neglect was twofold: Lack of fast and conveniently accessible computers was a serious hindrance; and there was then a general preference for analytical theory (using paper and pencil) among scientists, in comparison with numerical work.

In contrast, now access to efficient computers is easily available, every smartphone has orders of magnitude more computer power than the mainframe machines available in 1987 at university computer centers. Simulations now constitute a very large fraction of all theoretical work that is published; and if analytical work involves approximations, it is now considered very natural that the accuracy of these approximations is tested by appropriate simulations!

Despite the fact that computer simulation today is such an important tool of research for theoretical physics, a corresponding role of simulations in the teaching of theoretical physics to some extent is still lacking. Mostly, the courses follow the traditional textbooks where the subfields of theoretical physics, such as mechanics, electrodynamics, quantum mechanics, and statistical mechanics, are developed analytically; simulation methods at best are the subject of a small specialized course. This "wallflower existence" of simulations in a teaching context does by no means match the importance of simulation methods, and thus, the present book fills a distressing gap and satisfies an urgent need.

The present book is suitable both as a textbook for a specialized course and for use by students or researchers for self-study. The book tries to find a compromise between teaching the theoretical foundations of Monte Carlo simulation approaches and the "learning-by-doing" concept through exercises and examples that the student can work out. This strategy is clearly visible in the Chaps. 2 and 3 of this book that are its basic core. There the fundamental theoretical aspects are explained in rather simple terms, and a practical guide is developed, which enables the reader to

do original work of his own. This was basically the content of the first edition of this book in 1988.

Since then, the methodology of Monte Carlo simulation has seen a fantastic development. In the more recent methodological additions to the technique, the details are getting more and more specialized and complicated. Realizing the need to outline this methodological progress in three further chapters that were successively added in the further editions, we found it necessary to give a more condensed and somewhat abstract presentation. It was felt that the reader, who has worked through the more pedagogical first part of the book, should be able to use this additional material, guiding the reader also to the appropriate original literature. In this way, the total length of the book did not run out of hand. The present sixth edition contains two additional chapters that were not present in the earlier editions: Chap. 7 "Rejection Free Monte Carlo" and Chap. 8 "Finite Size Scaling Tools for the Study of Interfacial Phenomena and Wetting". While Chap. 7 thus focuses the foundation and on new algorithms that have been developed in the last decades, Chap. 8 focuses on the proper analysis of Monte Carlo results for spatially inhomogeneous systems. This topic is of increasing importance for surface science and nanoscience. As a side remark, we mention that the novel "event chain algorithm" solved (2011) the long-standing problem of clarifying the equation of state for a fluid of hard disks. This was the problem already addressed (but by no means solved!) in the very first Monte Carlo study in statistical physics by Metropolis et al. (1953).

Since simulation software can nowadays be downloaded simply from the Internet, carrying out simulations is as easy as never before; practically anyone can now try a "simulation study." However, if this is done without a proper understanding of the underlying methods, the likelihood of producing low quality results of very little usefulness is very great. The present book guides the newcomer to avoid such misuse of Monte Carlo simulations, giving the necessary background on the methods, as well as their intrinsic limitations (finite size effects, judgement errors due to insufficient length of the runs, lack of self-averaging, etc.). Working with this book hence will hopefully greatly help to acquire the necessary experience for doing sound simulation work.

Mainz, Germany — Kurt Binder
Heidelberg, Germany — Dieter W. Heermann

Preface to the Fifth Edition

The material presented in this book was born out of a series of lectures at a Summer School held at Figueira da Foz (Portugal) in 1987. Since then, the field of computational physics has seen an enormous growth and stormy development.

Many new applications and application areas have been found. In the 1980s, we could not foresee this but hoped that the Monte Carlo method would find such widespread acceptance. We were thus very glad to bring the work forward to a second edition correcting some misprints. Since then and over the years and editions of this book, many chapters have been added accounting for the development of new methods and algorithms. However, the basics have remained stable over the years and still serve as an entry point for researchers who would like to apply the Monte Carlo method and perhaps want to develop new ideas. Appending these basics with chapters on newly developed methods has evolved this book a bit into the direction of a textbook giving an introduction and at the same time covering a very broad spectrum. The first part of the book explains the theoretical foundations of the Monte Carlo method as applied to statistical physics. Chapter 3 guides the reader to practical work by formulating simple exercises and giving hints to solve them. Hence, it is a kind of "primer" for the beginner, who can learn the technique by working through these two chapters in a few weeks of intense study. Alternatively, this material can be used as text for a short course in university teaching covering in one term. The following chapters describe some more sophisticated and advanced techniques, e.g., Chap. 4 describes cluster algorithms and reweighting techniques, Chap. 5 describes the basic aspects of quantum Monte Carlo methods, and Chap. 6 (newly added to the fifth edition) describes recent developments in the last decade, such as "expanded ensemble" methods to sample the energy density of states, e.g., the Wang–Landau algorithm, as well as methods to sample rare events, such as "transition path sampling." These chapters then should be useful even for the more experienced practitioner. However, no attempt is made to cover all existing applications of Monte Carlo methods to statistical physics in an encyclopedic style—such an attempt would make this book almost unreadable and unhandy. While the "classic" applications of Monte Carlo methods in the 1970s and 1980s of the last century now are simple examples that a student

can work out on his laptop as an exercise, this is not true for the recent developments described in the last chapter, of course, which often need heavy investment of computer time. Hence, no attempt could as yet be made to enrich the last chapters with exercises as well.

We are very grateful for the many comments, suggestions, and the pointing out of misprints that have been brought to our attention. We would like to thank the many colleagues with whom we had the pleasure to engage with into discussions and that in some way or the other have shaped our thinking and thus have indirectly influenced this work.

Mainz, Germany — Kurt Binder
Heidelberg, Germany — Dieter W. Heermann
July 2010

Preface to the Fourth Edition

At the beginning of the new millennium, computer simulation is a well-established method of doing physics research. By Monte Carlo study of models that are intractable by analytical methods, one closes important gaps in our understanding of physical reality. "Computer experiments" can be performed where one switches on interactions at will (or switches them off), and one can "measure" response functions inaccessible by experiment, one can work in reduced dimensionality ($d = 1$, $d = 2$), or one can explore higher-dimensional worlds. These are just a few examples out of many, on how one can get insight by going beyond experiments. A valuable advantage also is the possibility of recognizing important aspects of a problem by visualizing degrees of freedom of a complex many-body system in any desired detail!

These comments should suffice to explain why the simulational approach in physics becomes still more popular, and the number of research papers alone that use it certainly is of the same order as research papers containing experimental work only or current analytical calculations. However, there still is a strange mismatch between the strong role of simulations in physics research, and the relatively small part that is devoted to simulation in the teaching of physics. The present book thus plays a key role, because it contributes significantly to closing this gap. Students with a little background in statistical thermodynamics can use this book to learn how to do simulations, guided using program simulations on classical problems of statistical physics, like the Ising model or other spin models, percolation, the Lennard–Jones fluid, etc. The combination of coherent chapters presenting all the essentials of the techniques of both the generation of simulation "data" and their analysis with a multitude of exercises of widely varying difficulty provides useful material, indispensable for the beginner, but containing facets also useful for the expert.

This concept applied also to previous editions and has proven successful and useful. Nevertheless, the present edition includes not only significant updates to the chapters contained in the earlier editions, but also contains a rich new chapter where an introduction to quantum Monte Carlo methods is provided. This is a topic which

steadily gains more importance, and hence, including it should significantly improve the usefulness of the present book.

Again, it is a great pleasure to thank many colleagues for suggestions, as well as our own students for their questions—all these interactions have helped to improve the presentation of material in this book.

Mainz, Germany — Kurt Binder
Heidelberg, Germany — Dieter W. Heermann
May 2002

Preface to the Third Edition

The last ten years have seen an explosive growth in the computer power available to scientists. Simulations that needed access to big mainframe computers in the past are now feasible on the workstation or powerful personal computer available on everybody's desk. This ease with which physicists (and scientists in neighboring areas such as chemistry, biology, economic science) can carry out simulations of their own has caused a true scientific revolution, and thus, simulational approaches are extremely widespread.

However, teaching simulation methods in physics is still a somewhat neglected field at many universities. Although there is plenty of literature describing advanced applications (the old dream of predicting materials properties from known interactions between atoms or molecules is now reality in many cases!), there is still a lack of textbooks from which the interested student can learn the technique of Monte Carlo simulations and their proper analysis step by step.

Thus, the present book still fulfills a need and continues to be useful for students who wish to bridge gaps in their university education in a "do-it-yourself" basis and for university staff who can use it for courses. Also researchers in academia and industry who have recognized the need to catch up with these important developments will find this book invaluable.

This third edition differs from the first in two important respects: Printing errors have been eliminated, unclear formulations have been replaced by better ones, and so on. We are most indebted to Professor Kecheng Qin (Physics Department, Univ. Beijing) who translated the first edition into Chinese and on that occasion very efficiently helped us to track down all these minor inconsistencies. We have also added an entire new chapter "Some Important Recent Developments of the Monte Carlo Methodology," which describes technical breakthroughs such as cluster algorithms and histogram reweighting, which became established after the first edition was published and are now commonly used by many Monte Carlo practitioners. The many references (far more than 100) in this chapter will make this book

useful for the experienced researcher as well as the new student, who is encouraged to apply these techniques when working through the exercises in Chap. 3.

Finally, we wish to thank many colleagues for fruitful interactions, which have helped to improve this book.

Mainz, Germany — Kurt Binder
Heidelberg, Germany — Dieter W. Heermann
June 1997

Preface to the Earlier Editions

When learning very formal material, one comes to a stage where one thinks one has understood the material. Confronted with a "real-life" problem, the passivity of this understanding sometimes becomes painfully clear. To be able to solve the problem, ideas, methods, etc., need to be ready at hand. They must be mastered (become active knowledge) in order to employ them successfully. Starting from this idea, the leitmotif, or aim, of this book has been to close this gap as much as possible.

How can this be done? The material presented here was born out of a series of lectures at the Summer School held at Figueira da Foz (Portugal) in 1987. The series of lectures was split into two concurrent parts. In one part, the "formal material" was presented. Since the background of those attending varied widely, the presentation of the formal material was kept as pedagogic as possible.

In the formal part, the general ideas behind the Monte Carlo method were developed. The Monte Carlo method has now found widespread application in many branches of science such as physics, chemistry, and biology. Because of this, the scope of the lectures had to be narrowed down. We could not give a complete account and restricted the treatment to the application of the Monte Carlo method to the physics of phase transitions. Here particular emphasis is placed on finite size effects.

The more "informal" part of the lectures concentrated on the practical side. In a step-by-step fashion, those who attended the lectures were led from "easy" applications to more advanced algorithms. In this part, we truly tried to give life to the ideas and concepts. We hope that in this book, we have captured the spirit of the Summer School. There, the gap mentioned before narrowed, because many actively participated in both parts.

From the above, it is clear that the material on the Monte Carlo method presented in this book can be of use to many scientists. It can be used for an advanced undergraduate or graduate course. In fact, a draft of this book has been used for a course held at the University of Mainz. Not only do we present the algorithms in great depth, we also encourage the reader to actively participate by setting many problems to be worked out by the reader.

Also for researchers and scientists using the Monte Carlo method, this book contains material which may be of importance for their research. We treat, for example, the problem of statistical errors of a Monte Carlo estimate of a quantity. Consideration is also given to the problem of self-averaging.

We would like to thank first of all K. Kremer and D.P. Landau. Without their continuing collaboration and constructive criticism, this book would not have its present form. Thanks are also due to the students of the condensed matter theory group at the University of Mainz for their participation and critical reading of the manuscript. Special thanks go to M. DeMeo for running some of the programs.

Mainz, Germany
May 1988

Kurt Binder
Dieter W. Heermann

Contents

Chapter 1
Introduction: Purpose and Scope of This Volume, and Some General Comments

In recent years the method of computer simulation has started something like a revolution of science: the old division of physics (as well as chemistry, biology, etc.) into experimental and theoretical branches is no longer really complete. Rather, computer simulation has become a *third branch* complementary to the first two traditional approaches.

What, then, is the specific significance of computer simulation or computer experiments? The answer is simply that computer simulation yields *exact* information (apart from statistical errors, but these can be made as small as desired, at least in principle) on *model* systems which are precisely characterized. (For problems in statistical physics this means that parameters describing the Hamiltonian are known explicitly and exhaustively.)

In contrast, the information provided by analytic theory is exact only in rather rare cases, while in most other cases uncontrolled approximations are required. For example, statistical physics problems which are solvable for a three-dimensional geometry are idealized limiting cases such as ideal gases or ideal solutions, coupled harmonic oscillators, etc. The statistical mechanics of even very simple models, such as the three-dimensional Ising model, cannot be solved exactly, and much less is known about models with realistic potentials between the atomic degrees of freedom. Thus computer simulations are often designed to check the accuracy of some approximation made in the analytical treatment of a model.

Similarly, the information provided by experiment is almost never precisely characterized in the sense that the effective Hamiltonian of a given experimental sample is precisely known. Sometimes it is even controversial whether some experimentally observed phenomenon is intrinsic or due to some unknown impurity effects – remember that the chemical constitution of an experimental sample is known only approximately anyway. These are just a few examples from which it is clear that the comparison between analytic theory and experiment does not always lead to conclusive answers, and simulations are needed to bridge this gap. Thus, a direct comparison between a simulation of a model and experiment is not hampered by

K. Binder and D. W. Heermann, *Monte Carlo Simulation in Statistical Physics*,
Graduate Texts in Physics, https://doi.org/10.1007/978-3-030-10758-1_1

inaccurate approximations, as are often inevitable in analytic theory, and hence may indicate more conclusively whether the model faithfully represents the real system or not.

Of course, this is by no means the only reason why computer simulations are attractive. It should be noted that simulations provide information on model systems which is arbitrarily detailed, and whatever quantity the researcher may consider useful he may attempt to sample from the simulation. For example, scattering techniques applied to real systems usually yield information on two-particle correlation functions, but it is very difficult to obtain direct experimental information on triplet correlations or even higher-order correlations. In contrast, simulations can yield such higher-order correlations readily, at least in principle. And while the experimenter may change the temperature and pressure of his sample, he cannot as easily assess the effect of varying the interatomic potential. But arbitrary variations of interatomic potentials do not constitute a major difficulty for a computer simulation in any way. It is now quite clear that the method of computer simulation is of interest in its own right; it is a valid scientific approach to understanding the laws of nature, instructive to its practitioners in a way that is complementary to theory or experiment.

In this situation, it is no surprise that there is a true explosion of the literature on the subject. Many researchers who have previously been doing research in theoretical physics (or theoretical chemistry, biology, etc.) start doing simulations, as well as some experimentalists. And, last but not least, many students who do not have any other research experience are attracted to the field of computer simulation immediately.

This great interest, however, encounters a serious difficulty: at this point, there is hardly any teaching of simulation methods at universities, and there is even a lack of systematic textbooks from which the newcomer to the field could easily learn to become an experienced practitioner. Although one of the authors (K.B.) of the present book has edited several books which collect many applications of the Monte Carlo computer simulation method in statistical physics, these books do not have the character of textbooks from which one can easily learn a new field. The other author (D.W.H.) has written a more pedagogic account of computer simulation methods in general; however, due to its generality it cannot go into very great detail as far as the Monte Carlo investigation of phase transitions and related problems (percolation, random walks, polymers, growth phenomena, etc.) is concerned. Similar reservations apply to other techniques (such as the molecular dynamics method) or the techniques have other limitations. Thus the art of Monte Carlo simulation so far is predominantly being learned and spread in two ways, namely, either by the tedious comparative study of many original papers dating back over several decades, or by private communications from experienced practitioners.

The purpose of the present book is to fill this gap, at least partially. Thus from the outset we restrict the scope of the book to *one* method of computer simulation, the Monte Carlo method, rather than trying to cover the whole field. This restriction in scope has several motivations: first of all, the expertise of the authors is mostly connected with this field: second, by this restriction it is realistic to use this book as a textbook for a two hour per week university course on computer simulation during

one university term. Alternatively, it is suitable for use as a text for a two-week workshop on computer simulation, where the student may practice every day during this two-week period, and thus learn the Monte Carlo method in a compact intensive course. Finally, for a student or researcher who tries to work through this book just by himself, the task still seems manageable!

Unlike previous literature on Monte Carlo simulation, the present book gives *equal weight* to the *theoretical foundations* of the method (including the analysis of the results) and to *practical work* with the method. Performing computer experiments must be learned, just as the experimentalist learns to plan and set up experiments with real systems and evaluate the data gained from them by attending practical courses. This need for practical work in order to learn to carry out such computer experiments has been encountered again and again both by the authors of this book and by many of their colleagues. In fact, preliminary unpublished notes for the present book have been used rather successfully for a workshop on computer simulation held at Figueira da Foz, Portugal, in September 1987, and at various courses held at the University of Mainz. Thus practical experience in teaching Monte Carlo methods to students was a major factor in determining the content of this book. It has been our experience that background knowledge of a programming language such as PASCAL can always be assumed, as well as some knowledge of statistical mechanics, including the basic principle of phase transitions. If the reader is not yet familiar with concepts such as critical exponents and the scaling relations among them and models such as the Ising model, percolation, etc., he can easily find various texts where these concepts are described clearly (we refer to some of these in this book). Thus there is no need to repeat these basic concepts.

However, in using the present book it is crucial to use the theoretical part (Chap. 2 in this book) together with the guide to practical work (Chap. 3). These chapters both deal with the same subjects (simple sampling, random and self-avoiding walks, percolation, the Ising model, etc.) but from somewhat different points of view. In the first part, concepts for the numerical treatment of these problems were introduced and justified. In the second part, these concepts are applied to problems, and *active participation* by the reader (e.g., by working on these problems on a personal computer) is required in order to understand the concepts more deeply.

A particularly suitable way of doing so is the form of a workshop where this text is used as the instruction manual. A solution to a problem is presented and immediately tried out, and the method for solving the problem, the algorithm, is improved upon. Of course, a workshop works best if there is interaction between the students and the teacher and among the students. There is a component of feedback, from which everybody in the workshop benefits. In the form of a written text a workshop is somewhat less efficient. Nevertheless, we have structured the text such that *some form of interaction with the text, other than passive reading, is possible and necessary.*

The aim is to present enough material so that one can start to develop algorithms for other problems based on the concepts presented here. To achieve this goal it is necessary to work through the entire material. Thus this workshop (Chap. 3) is a single unit. A second goal of Chap. 3 is to present methods of data analysis and

to enable the reader to become familiar with how they are applied. Again, active participation is requested.

With the concept used for this book with two chapters which are strongly correlated with each other, some redundancy is inevitable and even necessary for the sake of clarity and coherence of presentation. In fact, the scientific background of all the methods discussed in this book has been presented elsewhere in the literature: what is new and radically different from previous work is the introductory character which smoothly leads the student to a lot of practical work and experience with the method. For this pedagogic goal slight redundancies are even desirable. We have deliberately selected very simple problems of statistical physics, such as random and self-avoiding walk, percolation and the Ising model, for which all concepts and methods can be explained and demonstrated comparatively easily, and do not treat more complicated problems such as fluids with realistic potentials, spin glasses and other disordered materials, quantum-mechanical Monte Carlo methods, or problems in lattice gauge theory, in this part of the book. In our opinion, the reader will be able to move on to such problems using the other books which exist already on the Monte Carlo method, after he has worked through the present text. We deal with the characteristic features of thermal averaging for lattice problems with discrete degrees of freedom (Ising model, Potts model, etc.) as well as continuous ones (Heisenberg and XY magnets, ϕ^4 model, etc.) in some depth, while off-lattice problems such as simple fluids, are mentioned only briefly. Particular attention is paid to understanding the limitations of the method (effects due to finite size and boundary conditions, finite observation time effects, the question of self-averaging), and what one does to overcome these limitations: for example, finite-size effects at second-order phase transitions as well as at first-order phase transitions can be used as a valuable tool for studying the bulk properties of the system, if the appropriate finite-size scaling theory is invoked. The dynamic interpretation of the Monte Carlo importance sampling is discussed as well. It is shown that although on the one hand an unwanted slowing down of convergence is implied, particularly near critical points (critical slowing down) or in glassy systems, on the other hand the Monte Carlo method becomes a unique tool for the study of the kinetics of stochastic models.

When the reader has worked through Chaps. 2 and 3, he should have acquired enough familiarity with the basic tools of Monte Carlo methods, allowing him to proceed to the more advanced simulation methods. These are dealt with in Chaps. 4–8 in a more condensed form only, to avoid an excessive length of this book.

Part I
Learning the Basic Tools

Chapter 2
Theoretical Foundations of the Monte Carlo Method and Its Applications in Statistical Physics

In this chapter we first introduce the basic concepts of Monte Carlo sampling, give some details on how Monte Carlo programs need to be organized, and then proceed to the interpretation and analysis of Monte Carlo results.

2.1 Simple Sampling Versus Importance Sampling

2.1.1 Models

Statistical physics deals with systems with many degrees of freedom. A typical problem posed by statistical physics is to compute "average" macroscopic observables of a system for which the Hamiltonian is assumed to be known. For instance, let us consider magnetic systems: if a ferromagnet has very strong uniaxial anisotropy we may describe it by the Ising model, where N spins S_i interact as

$$\mathcal{H}_{\text{Ising}} = -J \sum_{\langle i,j \rangle} S_i S_j - H \sum_i S_i, \quad S_i = \pm 1, \tag{2.1}$$

where the spin S_i at lattice site i can point up or down along the "easy axis", the exchange energy J is restricted in (2.1) to nearest neighbors, and H is a magnetic field (the term $-H \sum_i S_i$ describing the Zeeman energy of the system). Other cases occur, however, if the ferromagnet has planar anisotropy (the spin being restricted to lie in the xy plane: XY model) or is fully isotropic (Heisenberg model):

K. Binder and D. W. Heermann, *Monte Carlo Simulation in Statistical Physics*,
Graduate Texts in Physics, https://doi.org/10.1007/978-3-030-10758-1_2

$$\mathcal{H}_{XY} = -J \sum_{\langle i,j \rangle} (S_i^x S_j^x + S_i^y S_j^y) - H_x \sum_i S_i^x, \tag{2.2}$$

$$(S_i^x)^2 + (S_i^y)^2 = 1,$$

$$\mathcal{H}_{\text{Heisenberg}} = -J \sum_{\langle i,j \rangle} (\boldsymbol{S}_i \cdot \boldsymbol{S}_j) - H_z \sum_i S_i^z, \tag{2.3}$$

$$(S_i^x)^2 + (S_i^y)^2 + (S_i^z)^2 = 1.$$

Of course, the large variety of real materials that the experimentalist can prepare in his laboratory creates interest in many variants of these models: instead of spin quantum number $S = \frac{1}{2}$, implied in (2.1), or $S \to \infty$, implied in (2.2) and (2.3), we may wish to consider general spin quantum numbers; instead of exchange between nearest neighbors only, we may wish to include exchange energies between next nearest neighbors, third nearest neighbors, etc.; instead of the full isotropy in (2.3), there may be a need to add a uniaxial or planar anisotropy term to it; instead of uniform exchange J and uniform field H in (2.1), it may be appropriate to work with random exchange constants J_{ij} and random fields H_i, to model some frozen-in random disorder in the system. Thus, magnetic solids already provide us with an incredible wealth of model Hamiltonians, for which (2.1)–(2.3) just provide prototype examples, and this wealth of models is only a small part of the broad spectrum of applications provided by condensed matter physics.

One task of statistical physics is to compute from the model Hamiltonian $\mathcal{H}$ the desired average properties, e.g., the average energy E or average magnetization $\boldsymbol{M}$ per degree of freedom,

$$E = \langle \mathcal{H} \rangle_T / N, \quad \mathbf{M} = \left\langle \sum_i \vec{S}_i \right\rangle_T / N. \tag{2.4}$$

Here the thermal average of any observable $A(\boldsymbol{x})[A = \mathcal{H}, \sum_i \vec{S}_i$, etc., and the vector $\boldsymbol{x}$ in phase space stands symbolically for the set of variables describing the considered degree of freedom, e.g., $\boldsymbol{x} = (S_1, S_2, \ldots, S_N)$ for (2.1) $\boldsymbol{x} = (\vec{S}_1, \vec{S}_2, \ldots, \vec{S}_N)$ for (2.3)] is defined in the canonical ensemble

$$\langle A(\boldsymbol{x}) \rangle_T = \frac{1}{Z} \int \mathrm{d}\boldsymbol{x} \exp\left[-\mathcal{H}(\boldsymbol{x})/k_\mathrm{B}T\right] A(\boldsymbol{x}), \tag{2.5}$$

$$Z = \int \mathrm{d}\boldsymbol{x} \exp\left[-\mathcal{H}(\boldsymbol{x})/k_\mathrm{B}T\right].$$

It is appropriate to call these classes of problems "statistical physics" because the normalized Boltzmann factor

$$p(\boldsymbol{x}) = \frac{1}{Z} \exp\left[-\mathcal{H}(\boldsymbol{x})/k_\mathrm{B}T\right] \tag{2.6}$$

plays the role of a probability density describing the statistical weight with which the configuration $\boldsymbol{x}$ occurs in thermal equilibrium.

Now although (2.6) gives a formally exact description of the probability distribution $p(\boldsymbol{x})$, we are still in trouble: we are neither interested in such detailed information (in our examples $\boldsymbol{x}$ stands for a set containing the N spin degrees of freedom), nor is it possible to carry out the integrations in this high-dimensional space [(2.4) and (2.5)] in the general case.

2.1.2 Simple Sampling

The Monte Carlo method in equilibrium statistical mechanics starts from the idea of approximating the exact equation (2.5), where one integrates over all states $\{\boldsymbol{x}\}$ with their proper weights $p(\boldsymbol{x})$, by a summation using only a characteristic subset of phase space points $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_M\}$ which are used as a statistical sample. Clearly, if one considers the limit $M \to \infty$, the discrete sum

$$\overline{A(\boldsymbol{x})} = \frac{\sum_{l=1}^{M} \exp\left[-\mathcal{H}(\boldsymbol{x}_l)/k_{\mathrm{B}}T\right] A(\boldsymbol{x}_l)}{\sum_{l=1}^{M} \exp\left[-\mathcal{H}(\boldsymbol{x}_l)/k_{\mathrm{B}}T\right]} \tag{2.7}$$

must approximate (2.5), just as in numerical integration routines integrals are replaced by sums [for discrete degrees of freedom, such as the Ising problem, $\int \mathrm{d}\boldsymbol{x}$ in (2.5) already stands for a discrete summation over all the 2^N states $\boldsymbol{x} = (S_1, \ldots, S_N)$, of course, but in (2.7) we then wish to work with a small subset of these states only, $M \ll 2^N$]. But, unlike in standard routines to solve one-dimensional integrals $\int f(x)\mathrm{d}x$, where $f(x)$ is a function of one real variable x only, instead of a high-dimensional vector $\boldsymbol{x}$, it makes no sense to choose the points $\boldsymbol{x}_l$ according to a regular grid, rather we have to choose the points $\boldsymbol{x}_l$ at random. In order to appreciate this point in detail, let us consider the XY model defined in (2.2) as an example. Because $(S_i^x)^2 + (S_i^y)^2 = 1$ for each site i, it is convenient to write $S_i^x = \cos\varphi_i$, $S_i^y = \sin\varphi_i$ and take the angle $\varphi_i (0 \le \varphi_i < 2\pi)$ as a variable to characterize the degrees of freedom. Then $\int \mathrm{d}\boldsymbol{x}$ simply means $\prod \int_0^{2\pi} \mathrm{d}\varphi_i$. Let us now introduce a regular grid, defined by $\varphi_i^{\gamma} = (\gamma_i/p)2\pi$, with $\gamma_i = 1, 2, \ldots, p$, where p is some integer characterizing the grid. Obviously the total number of points to be used in this grid is p^N, which is very large for large N, impossible to use in practice even if p is rather small. Apart from this difficulty, even if we were able to work with a reasonably large value for p, we would still have the problem that almost all points were located on the surface of the integration hypercube and almost none in its interior. Since in any lattice direction of the hypercube there are p points of the grid, $p - 2$ being in the cube interior, the total fraction of points in the interior is

$$[(p-2)/p]^N = (1-2/p)^N$$
$$= \exp\left[N \log\left(1-\frac{2}{p}\right)\right] \underset{p \text{ large}}{\approx} \exp\left[-\frac{2N}{p}\right] \underset{N\to\infty}{\longrightarrow} 0.$$

A much better, i.e., *uniform*, distribution of grid points is achieved if we choose the points x_l *at random*, utilizing "*pseudo-random numbers*" produced by a "*random number generator*" built into the computer. This use of random numbers has given this game its name! In fact, the method described thus far by (2.7) is indeed a variant of Monte Carlo methods, namely the *simple sampling* Monte Carlo method.

2.1.3 Random Walks and Self-avoiding Walks

As an example of problems for which the simple sampling technique has actually been and is still used, we mention the study of self-avoiding walks (SAWs) on lattices, see, e.g., [2.1]. These self-avoiding walks are used for modelling the large-scale properties of long flexible macromolecules in solution [2.2–2.4]. Since it is rather instructive to discuss both the advantages and the disadvantages of studying such random-walk-type problems with simple sampling Monte Carlo methods, we give a brief digression on this subject in the following.

Figure 2.1 shows various types of random walks on the square lattice. There are four different types of vectors $\boldsymbol{v}(k)$ connecting a site to its nearest neighbor on the lattice (the lattice spacing is taken to be unity)

$$\boldsymbol{v}(1) = (1,0), \quad \boldsymbol{v}(2) = (0,1), \quad \boldsymbol{v}(3) = (-1,0), \quad \boldsymbol{v}(4) = (0,-1). \tag{2.8}$$

An algorithm which generates simple (unrestricted) random walks of N steps now proceeds as follows:

Algorithm 2.1 Random walks

i) $\boldsymbol{r}_0 = 0$ (origin of coordinate system) and put $k = 0$
ii) Choose a random integer ν_k between 1 and 4
iii) Replace k by $k+1$ and put $\boldsymbol{r}_k = \boldsymbol{r}_{k-1} + \boldsymbol{v}(\nu_{k-1})$
iv) If $k = N$ put $\boldsymbol{r}_k = \boldsymbol{R}$ (end to end distance of the walk); else return to step (ii). (2.9)

An example of a walk with $N = 22$ generated by this procedure is shown in Fig. 2.1a; the generation of random walks will be studied in more detail in Sect. 3.2.1. At this point we only note that for a lattice of coordination number z the total number Z_N of all such (different) random walks (RWs) is simply

$$Z_N^{\mathrm{RW}} = z^N. \tag{2.10}$$

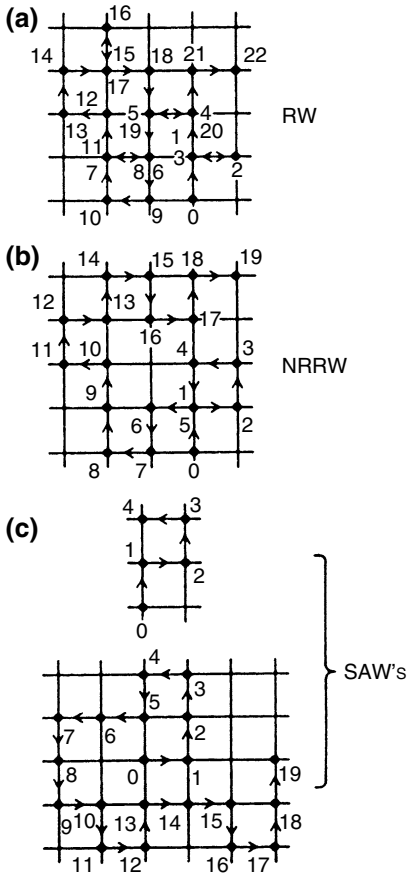


Fig. 2.1 An unrestricted random walk (RW) of 22 steps on the square lattice (**a**), a nonreversal random walk (NRRW) (**b**), and two self-avoiding walks (SAWs) (**c**). Sites are numbered in the order that they are visited. Bonds with *arrows* are selected consecutively by means of random numbers. The *dots* on the lattice sites then represent the monomers of the polymer chain

If the random walk is taken as a model for a polymer chain, Z_N is just the polymer partition function. (In the absence of any interactions all chain configurations have exactly the same statistical weight.)

While Algorithm 2.1 may be a reasonable model for hopping conduction in solids or other diffusion processes on ideal lattices, it is not a good model for polymers in solution, not just because of the unrealistic features of using a lattice structure to model the conformations of a macromolecule, but in particular because the excluded volume interaction is ignored. Unlike real polymers, the random walk in Fig. 2.1a intersects itself and folds back on itself. The latter feature is eliminated by defining the nonreversal random walk (NRRW) for which immediate reversals are forbidden. We can define an algorithm for this NRRW by introducing a sort of "periodic" boundary condition for the vectors $\boldsymbol{v}(\nu_k)$ by defining

$$\boldsymbol{v}(\nu \pm 4) = \boldsymbol{v}(\nu) \tag{2.11}$$

and modifying step (ii) of (2.9) for $k > 1$ by introducing a one-step memory:
(ii′): Choose a random number out of the triplet

$$\{\nu_{k-1} - 1, \nu_{k-1}, \nu_{k-1} + 1\} \text{ and take it as } \nu_k. \tag{2.12}$$

An alternative realization of the NRRW would proceed as in (2.9) but would throw ν_k away if $\boldsymbol{v}(\nu_k) = \boldsymbol{v}(\nu_{k-1} + 2)$, using (2.11) if necessary, and iterating step (ii). In this way the same random numbers yielding the RW with $N = 22$ in Fig. 2.1a yield a NRRW with $N = 19$ in Fig. 2.1b. From (2.12) we realize that

$$Z_N^{\mathrm{NRRW}} = (z-1)^N. \tag{2.13}$$

At this point, the reader should work through Sect. 3.2.2 to gain insight into the NRRW from practical work. It is clear that the NRRW algorithm still leads to chain configurations which may intersect, which is a qualitatively unrealistic feature of real macromolecule configurations as mentioned above. We can change our algorithm in order to generate self-avoiding walks (SAWs) instead of NRRWs by adding in (2.9) between step (iii) and step (iv) the further condition:

(iii′) If $\boldsymbol{r}_k$ leads to a lattice site which has already been visited by this walk, stop the construction, return to (i) and start the construction all over again. (2.14)

It turns out that in this case the number of configurations no longer has the simple exponential form as in (2.10) or (2.13), instead, a simple form holds only in the asymptotic limit $N \to \infty$, and even then involves a power-law correction term,

$$Z_N^{\mathrm{SAW}} \underset{N\to\infty}{\longrightarrow} N^{\gamma-1} z_{\mathrm{eff}}^N, \quad z_{\mathrm{eff}} \le z - 1. \tag{2.15}$$

Note that γ is a critical exponent, and the effective coordination number z_{eff} in general is noninteger ($z_{\mathrm{eff}} = z - 1$ only for $z = 2$; the one-dimensional lattice). For *three*-dimensional lattices neither γ nor z_{eff} can be calculated by exact methods analytically; therefore Monte Carlo methods play an important role in estimating these quantities characterizing the SAW statistics.

Figure 2.1c shows that the same random numbers drawn for the construction of the RW or the NRRW example would terminate the SAW already at the fifth step, and hence according to (2.14) this trial to construct a SAW is unsuccessful, and another attempt needs to be made. Clearly, for large N most attempts will not be successful. We can estimate the fraction of successful attempts (i.e., the probability p_N that a NRRW of N steps is self-avoiding) simply by taking the ratio of the respective partition functions:

$$p_N = \frac{Z_N^{\mathrm{SAW}}}{Z_N^{\mathrm{NRRW}}} \underset{N\to\infty}{\longrightarrow} N^{\gamma-1}\left(\frac{z_{\mathrm{eff}}}{z-1}\right)^N$$
$$= \exp\left(-N\ln\frac{z-1}{z_{\mathrm{eff}}} + (\gamma-1)\ln N\right). \tag{2.16}$$

Thus, for large N the probability of succeeding in getting SAWs decreases exponentially fast with N; this inefficiency is called the *attrition problem*. Therefore, practical applications of this simple random sampling of SAWs are restricted to $N \leq 100$. See Fig. 2.2 for an example. In this example, a somewhat generalized problem is considered: in addition to the excluded volume interaction (that is, an infinitely high repulsive potential if two different monomers occupy the same site) an attractive energy $(-\varepsilon, \varepsilon > 0)$ is included if two monomers occupy nearest-neighbor sites on the lattice. It is then of interest to study the internal energy $\langle \mathcal{H} \rangle_T$ of the chain as well as the chain average linear dimensions (such as $\langle R^2 \rangle$) as a function of the reduced temperature $k_{\mathrm{B}}T/\varepsilon$, since for $N \to \infty$ there occurs at $T = \theta$ the so-called collapse transition of the chain [2.1, 2.3]: for $T > \theta$ the chain linear dimension is "swollen" in comparison to a simple random walk, for which simply

$$\langle R^2 \rangle_{T=\infty}^{\mathrm{RW}} = N. \tag{2.17}$$

Instead of (2.17) we have for self-avoiding walks a law involving another critical exponent (ν)

$$\langle R^2 \rangle_{T>\theta}^{\mathrm{SAW}} \propto N^{2\nu}, \quad \nu \approx 0.59, \quad N \to \infty. \tag{2.18}$$

Due to the attractive energy between the monomers, the chain "collapses" for $T < \theta$, i.e., it takes a rather compact configuration described by

$$\langle R^2 \rangle_{T<\theta}^{\mathrm{SAW}} \propto N^{2/3}, \quad N \to \infty, \tag{2.19}$$

while exactly for $T = \theta$ the chain configuration is asymptotically similar to that of an ideal (Gaussian) random walk,

$$\langle R^2 \rangle_{T=\theta}^{\mathrm{SAW}} \propto N, \quad N \to \infty. \tag{2.20}$$

Note that (2.17) is easily proven exactly, since from (2.9) we immediately conclude

$$\boldsymbol{R} = \sum_{k=1}^{N} \boldsymbol{v}(\nu_k),$$
$$\langle \boldsymbol{R}^2 \rangle_{T=\infty}^{\mathrm{RW}} = \sum_{k=1}^{N} \boldsymbol{v}^2(\nu_k) + \sum_{\substack{k,k' \\ (k\neq k')}} \langle \boldsymbol{v}(\nu_k) \cdot \boldsymbol{v}(\nu_{k'}) \rangle_{T=\infty}, \tag{2.21}$$

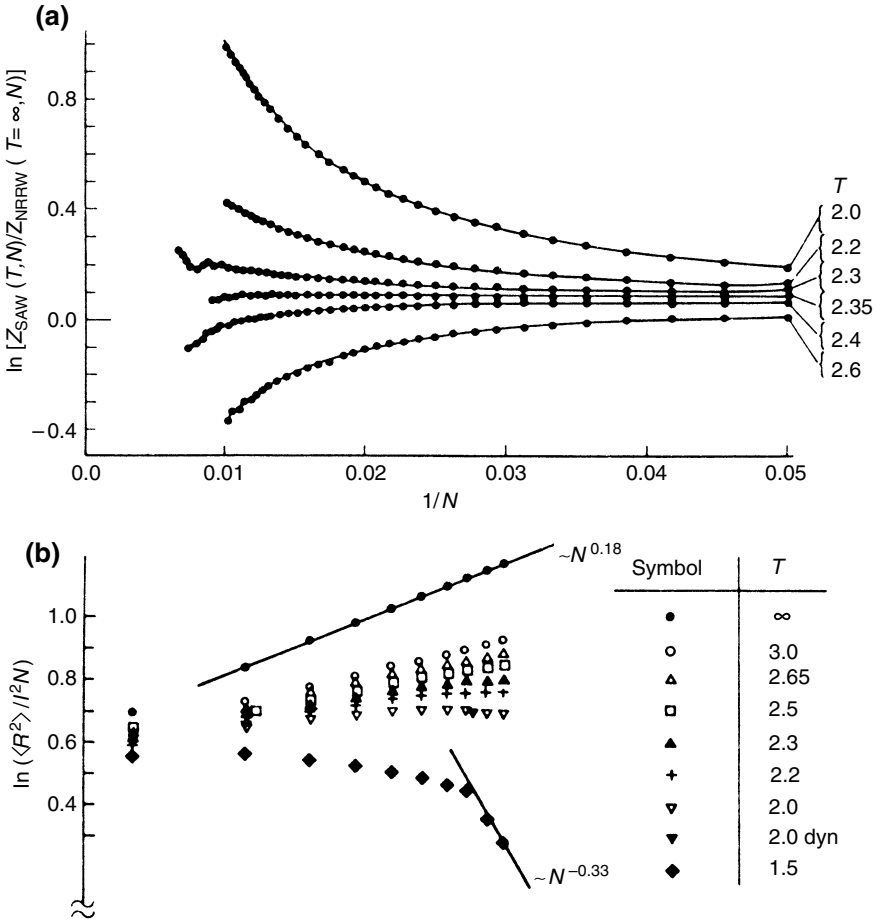


Fig. 2.2 **a** Simple-sampling data (with 2×10^6 samples for $N = 100$) for the free-energy difference $\ln[Z_{SAW}(T, N)/Z_{NRRW}(T=\infty, N)]$ plotted versus $1/N$, for various temperatures (measured in units of ε/k_B) and the tetrahedral lattice. **b** Log–log plot of $\langle R^2\rangle/l^2N$ against N, l being the bond length ($l = \sqrt{3}$) on the tetrahedral lattice. Data are obtained by simple sampling methods; at $T = 2$, $N = 80$ a result of dynamical simulations with motions of three and four bonds is included. (From [2.1])

noting that $\boldsymbol{v}^2(\nu_k) = 1$, from (2.8), while $\langle \boldsymbol{v}(\nu_k) \cdot \boldsymbol{v}(\nu_{k'})\rangle_{T=\infty} = 0$ since there are no correlations between different steps of the walk. On the other hand, in the case of (2.18) and (2.19) obviously there must be strong correlations even between very distant steps of the walk. Thus it is no surprise that it is possible neither to derive [(2.18) and (2.19)] analytically nor to locate the "theta temperature" $k_B\theta/\varepsilon$, where (2.20) holds, analytically, and thus Monte Carlo methods are useful to study a problem of this kind.

At this point it is useful to briefly discuss the practical implementation of this simple sampling method for SAWs (for more details, the reader should work through Sect. 3.2.3). First of all we have to keep track of the previous steps of the walk to make sure that it does not intersect itself. One could store all vectors $\boldsymbol{r}_{k'}$ previously generated in the construction of a walk, $0 \le k' \le k-2$, and check whether there exist a k' such that $\boldsymbol{r}_k = \boldsymbol{r}_{k'}$. Of course, such an algorithm would be very time consuming, and it is much more efficient to choose a finite lattice (e.g., of dimension $(2N+1) \times (2N+1)$ if we wish to study SAWs of N steps at the square lattice) where we introduce occupation variables c_i for each lattice site i and initially put $c_i = 0$ for all i. Now generating a walk we put $c_i = 1$ if it visits site i, and we generate an array labeling the lattice sites which have $c_i = 1$ in order in which they have been visited. Now the excluded volume condition is easily satisfied by checking at each step whether $c_i = 0$ still at the next site to be visited. After the termination of the construction (and possible analysis of the SAW configuration) we have to go through the array of sites which have been visited to replace $c_i = 1$ by $c_i = 0$ again and then we can start the next construction of the walk. Alternatively, one can increase a label by one and check if the next site to be visited is equal to the label.

Note also that for studying laws such as (2.15), (2.18)–(2.20) one wishes not only to study one particular value of N but a whole range of values for N. There is no need to repeat the procedure for different choices of N; instead we can put N to the largest value we wish to analyze and just sample in addition all the walks which are unsuccessful constructions in the sense of (2.16), i.e., they terminate at some value $N' < N$ since in the $(N'+1)$th step one would have violated the excluded volume restriction.

2.1.4 Thermal Averages by the Simple Sampling Method

How does temperature come into the game? By definition of the model, a configuration with n nearest-neighbor contacts (other than those along the contour of the chain itself) will have a Boltzmann weight factor proportional to $\exp(n\varepsilon/k_{\mathrm{B}}T)$. So we need to keep track of the number n in each configuration and generate the appropriate distribution function: the thermal averaging at any temperature T that one wishes to study can then be done afterwards! Specifically, the Monte Carlo sampling attempts to estimate the distribution $p_N(n) = Z_N^{\mathrm{SAW}}(n)/Z_N^{\mathrm{NRRW}}$, i.e., the (normalized) number of SAW configurations of N steps with n nearest-neighbor contacts, and $p_N(n, \boldsymbol{R}) = Z_N^{\mathrm{SAW}}(n, \boldsymbol{R})/Z_N^{\mathrm{NRRW}}$, the number of SAW configurations of N steps with n nearest-neighbor contacts and an end-to-end vector $\boldsymbol{R}$. Then the averages of interest can be expressed as follows:

$$\langle \mathcal{H} \rangle_T = \frac{-\varepsilon \sum_n n \exp(n\varepsilon/k_\mathrm{B}T)\, p_N(n)}{\sum_n \exp(n\varepsilon/k_\mathrm{B}T)\, p_N(n)}, \tag{2.22a}$$

$$\langle R^2 \rangle_T = \frac{\sum_{n,\boldsymbol{R}} R^2 \exp(n\varepsilon/k_\mathrm{B}T)\, p_N(n, \boldsymbol{R})}{\sum_n \exp(n\varepsilon/k_\mathrm{B}T)\, p_N(n)}. \tag{2.22b}$$

The specific heat C per bond of the chain can be obtained utilizing the fluctuation relation

$$\frac{C}{k_\mathrm{B}} = \frac{1}{N} \frac{\partial(\langle \mathcal{H} \rangle_T)}{\partial(k_\mathrm{B}T)} = \frac{\langle \mathcal{H}^2 \rangle_T - \langle \mathcal{H} \rangle^2}{N k_\mathrm{B}^2 T^2} \tag{2.23}$$

to find [note that (2.23) is easily verified from (2.22a) and the definition of C, but holds generally]

$$\frac{C}{k_\mathrm{B}} = \left(\frac{\varepsilon}{k_\mathrm{B}T}\right)^2 \frac{\left(\sum_n n^2 p_N(n) \mathrm{e}^{n\varepsilon/k_\mathrm{B}T} - \left[\sum_n n p_N(n) \mathrm{e}^{n\varepsilon/k_\mathrm{B}T} \right]^2 \right)}{N \sum_n p_N(n) \mathrm{e}^{n\varepsilon/k_\mathrm{B}T}}. \tag{2.24}$$

Figure 2.2 presents some examples found from a simple-sampling study of self-avoiding walks on the tetrahedral lattice [2.1].

2.1.5 Advantages and Limitations of Simple Sampling

Simple sampling of self-avoiding walks as described so far has two advantages: (1) In one simulation run we obtain information on the full range of values for chain length N up to some maximum length, and for a broad range of temperatures, (2) The individual configurations of the walks are statistically independent of each other, and therefore standard error analysis applies. Suppose M configurations of N-step walks have been generated successfully. Then $\langle R^2 \rangle_{T=\infty}$ (i.e., in the athermal case) is obtained as follows, applying (2.7):

$$\langle R^2 \rangle_{T=\infty} \approx \overline{R^2} = \frac{1}{M} \sum_{l=1}^{M} R_l^2, \tag{2.25}$$

$\boldsymbol{R}_l$ being the end-to-end distance of the lth configuration of the walk, and its error $\langle (\delta R^2)^2 \rangle_{T=\infty}$ is estimated as

$$\langle(\delta R^2)^2\rangle_{T=\infty} \approx \overline{(\delta R^2)^2} = \frac{1}{M(M-1)} \sum_{l=1}^{M} \left[R_l^4 - \overline{(R^2)}^2 \right]. \tag{2.26}$$

For the random walk or the nonreversal random walk it is easy to predict the expected relative error, making use of the Gaussian character of the probability distribution [2.2, 2.3]:

$$p_N(\boldsymbol{R}) \propto \exp\left[-dR^2/\left(2\langle R^2\rangle\right)\right], \quad d \text{ is the dimensionality;} \tag{2.27}$$

remember that

$$\begin{aligned} p_N(\boldsymbol{R}) &= p_N(x_1)\, p_N(x_2) \ldots p_N(x_d) \\ &\propto \exp\left(-\frac{x_1^2}{2\langle x_1^2\rangle}\right) \exp\left(-\frac{x_2^2}{2\langle x_2^2\rangle}\right) \ldots \exp\left(-\frac{x_d^2}{2\langle x_d^2\rangle}\right) \quad \text{and} \\ \langle x_1^2\rangle &= \langle x_2^2\rangle \ldots \langle x_d^2\rangle = \langle R^2\rangle/d \end{aligned}$$

in an isotropic system where $R^2 = x_1^2 + x_2^2 + \cdots + x_d^2$. From (2.27) one obtains

$$\begin{aligned} \langle R^4\rangle_{T=\infty} &= \frac{\displaystyle\int_0^\infty R^{d+3} \mathrm{d}R \exp\left(-\frac{dR^2}{2\langle R^2\rangle}\right)}{\displaystyle\int_0^\infty R^{d-1} \mathrm{d}R \exp\left(-\frac{dR^2}{2\langle R^2\rangle}\right)} \\ &= \left(\frac{2\langle R^2\rangle}{d}\right)^2 \frac{\Gamma(d/2+2)}{\Gamma(d/2)} = \langle R^2\rangle^2 \frac{d+2}{d}, \end{aligned} \tag{2.28}$$

and hence the relative error is estimated as

$$\frac{\overline{(\delta R^2)^2}}{\overline{(R^2)}^2} \cong \frac{1}{(M-1)} \frac{\left(\langle R^4\rangle_{T=\infty} - \langle R^2\rangle_{T=\infty}^2\right)}{\langle R^2\rangle_{T=\infty}^2} = \frac{1}{(M-1)} \frac{2}{d}. \tag{2.29}$$

This is a very simple example of *lack of self-averaging*. One learns in thermodynamics that in the thermodynamic limit $N \to \infty$ fluctuations die out, and the relative fluctuation of extensive thermodynamic variables A vanishes, $\langle(\delta A)^2\rangle/\langle A\rangle^2 \propto 1/N \to 0$. In a large system, therefore, one single observation A_l of a quantity A in equilibrium is rather close to the average $\langle A\rangle$ already: A_l will differ from $\langle A\rangle$ only by terms of order $1/\sqrt{N}$. This property is called *strong self-averaging* [2.5]. This is obviously not true for the quantities of interest in a random walk problem, such as $\langle R^2\rangle$. The reason that there is no self-averaging for $\langle R^2\rangle$ is that $\langle R^2\rangle$ is itself a fluctuation, in a thermodynamic sense, and the quantity to which the thermodynamic argument applies is $\langle \boldsymbol{R}\rangle$ which according to (2.21) can be written

$$\langle \boldsymbol{R} \rangle = \sum_{k=1}^{N} \langle \boldsymbol{v}(\nu_k) \rangle = N \langle \boldsymbol{v} \rangle, \tag{2.30}$$

where $\langle \boldsymbol{v} \rangle \equiv 0$, however. If we were to consider a "biased" random walk, however, in which one particular step orientation is chosen with higher probability than all other step orientations such that $\langle \boldsymbol{v} \rangle \neq 0$, we would have $\langle \boldsymbol{R} \rangle^2 = N^2 \langle \boldsymbol{v} \rangle^2 \gg \langle \boldsymbol{R}^2 \rangle - \langle \boldsymbol{R} \rangle^2 \propto N$, and we would have the standard thermodynamic relation $\langle (\delta \boldsymbol{R})^2 \rangle / \langle \boldsymbol{R} \rangle^2 \propto 1/N$.

For self-avoiding walks the distribution function $p_N(\boldsymbol{R})$ is not a simple Gaussian, but we have the same property that $\langle R^4 \rangle_{T=\infty} \propto \langle R^2 \rangle^2$ (but there is no longer a simple argument yielding the proportionality factor). It then again follows that the relative error is independent of N, i.e., there occurs lack of self-averaging.

Apart from these random-walk-type problems and other nonthermal problems such as percolation [2.6], which will be discussed in Sect. 2.3.1 in another context, simple random sampling techniques are not so useful for evaluating thermal averages such as (2.7). Consider for instance the case where in (2.7) $A(\boldsymbol{x}_l)$ is the Hamiltonian $\mathcal{H}(\boldsymbol{x}_l)$ itself. Now (2.23) implies that the relative fluctuation $\left(\langle \mathcal{H}^2 \rangle_T - \langle \mathcal{H} \rangle_T^2\right) / \langle \mathcal{H} \rangle_T^2 \propto 1/N$, which implies that the probability distribution $p(E)$ of the energy E per degree of freedom defined as

$$p(E) = \frac{1}{Z} \int \mathrm{d}\boldsymbol{x}\ \delta \left(\mathcal{H}(\boldsymbol{x}) - NE \right) \exp \left[-\mathcal{H}(\boldsymbol{x}) / k_\mathrm{B} T \right] \tag{2.31}$$

is very sharply peaked, since we can also write

$$\langle \mathcal{H} \rangle_T = N \int_{-\infty}^{+\infty} E p(E) \mathrm{d}E, \quad \langle \mathcal{H}^2 \rangle_T = N^2 \int_{-\infty}^{+\infty} E^2 p(E) \mathrm{d}E, \tag{2.32}$$

and hence $p(E)$ must have a peak of height $\sqrt{N}$ and width $1/\sqrt{N}$ near $E = \langle \mathcal{H} \rangle_T / N$. In fact, off second- or first-order phase transitions one can show that $p(E)$ is actually again Gaussian [2.6]

$$p(E) \propto \exp \left(-\frac{(E - \langle \mathcal{H} \rangle_T / N)^2}{2 C k_\mathrm{B} T^2} N \right). \tag{2.33}$$

Figure 2.3 shows a schematic sketch of this probability distribution $p(E)$. Now a simple sampling in fact means a generation of phase space points $\{\boldsymbol{x}\}$ with a probability distribution $p(E)$ peaked at $E = 0$ [for models such as (2.1)–(2.3), where $\mathcal{H}_{T=\infty} = 0$], again with a Gaussian distribution of width $1/\sqrt{N}$. Thus the probability of generating states with E near $\langle \mathcal{H} \rangle / N$ is exponentially small, if the states are generated by the simple random sampling technique. (As a consequence of this fact, (2.22) and (2.24) are of no practical use at temperatures $T < \theta$, apart from very small values of N.) So what is needed is a more efficient technique that samples the configurations $\boldsymbol{x}_l$ included in the average (2.7) not completely at random, but preferentially from that regime of phase space which is important at temperature T. Suppose we

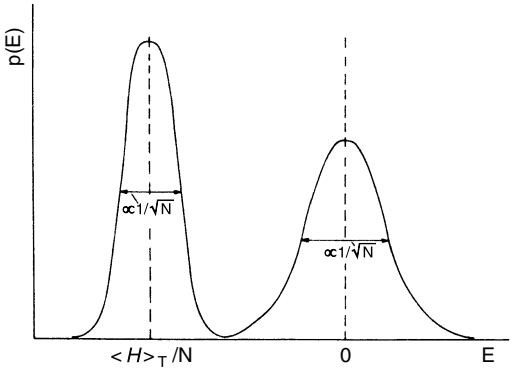


Fig. 2.3 Probability distribution $p(E)$ of the normalized energy E. The *right curve* (peaked at $E = 0$) is the probability distribution $p(E)$ generated by simple sampling of phase space points $\{\boldsymbol{x}\}$. (Note that for spin models such as the Ising model (2.1), states with positive E and with negative E are generated with the same probability.) The *left curve* is the actual probability distribution $p(E)$ occurring in the canonical ensemble at a *finite* temperature T. Note that these two probability distributions for large N overlap only in their wings

consider a process where the phase space points $\boldsymbol{x}_l$ are selected according to some probability $P(\boldsymbol{x}_l)$. Choosing this set $\{\boldsymbol{x}_l\}$ now for the estimation of a thermal average, (2.7) is replaced by

$$\overline{A(\boldsymbol{x})} = \frac{\sum_{l=1}^{M} \exp\left[-\mathcal{H}(\boldsymbol{x}_l)/k_\mathrm{B}T\right] A(\boldsymbol{x}_l)/P(\boldsymbol{x}_l)}{\sum_{l=1}^{M} \exp\left[-\mathcal{H}(\boldsymbol{x}_l)/k_\mathrm{B}T\right]/P(\boldsymbol{x}_l)}. \tag{2.34}$$

2.1.6 Importance Sampling

A simple and most natural choice for $P(\boldsymbol{x}_l)$ in (2.34) would be $P(\boldsymbol{x}_l) \propto \exp[-\mathcal{H}(\boldsymbol{x}_l)/k_\mathrm{B}T]$; then the Boltzmann factor cancels out altogether, and (2.34) reduces to a simple arithmetic average

$$\overline{A(\boldsymbol{x})} = \frac{1}{M} \sum_{l=1}^{M} A(\boldsymbol{x}_l). \tag{2.35}$$

The problem is, of course, to find a procedure which practically realizes this so-called importance sampling [2.7]. Metropolis et al. [2.7] advanced the idea not to choose the successive states $\{\boldsymbol{x}_l\}$ independently of each other, but to construct a Markov

process where each state $\boldsymbol{x}_{l+1}$ is constructed from a previous state $\boldsymbol{x}_l$ via a suitable transition probability $W(\boldsymbol{x}_l \to \boldsymbol{x}_{l+1})$. They pointed out that it is possible to choose the transition probability W such that in the limit $M \to \infty$ the distribution function $P(\boldsymbol{x}_l)$ of the states generated by this Markov process tends towards the equilibrium distribution

$$P_{\mathrm{eq}}(\boldsymbol{x}_l) = \frac{1}{Z} \exp\left(-\frac{\mathcal{H}(\boldsymbol{x}_l)}{k_{\mathrm{B}}T}\right) \tag{2.36}$$

as desired. A sufficient condition to achieve this is to impose the principle of detailed balance

$$P_{\mathrm{eq}}(\boldsymbol{x}_l) W(\boldsymbol{x}_l \to \boldsymbol{x}_{l'}) = P_{\mathrm{eq}}(\boldsymbol{x}_{l'}) W(\boldsymbol{x}_{l'} \to \boldsymbol{x}_l). \tag{2.37}$$

Equation (2.37) implies that the ratio of transition probabilities for a "move" $\boldsymbol{x}_l \to \boldsymbol{x}_{l'}$ and the inverse move $\boldsymbol{x}_{l'} \to \boldsymbol{x}_l$ depend only on the energy change $\delta\mathcal{H} = \mathcal{H}(\boldsymbol{x}_{l'}) - \mathcal{H}(\boldsymbol{x}_l)$,

$$\frac{W(\boldsymbol{x}_l \to \boldsymbol{x}_{l'})}{W(\boldsymbol{x}_{l'} \to \boldsymbol{x}_l)} = \exp\left(-\frac{\delta\mathcal{H}}{k_{\mathrm{B}}T}\right). \tag{2.38}$$

Equation (2.38) obviously does not specify $W(\boldsymbol{x}_l \to \boldsymbol{x}_{l'})$ uniquely, and some arbitrariness in the explicit choice of W remains. Two frequently used choices are [2.8, 2.9, 2.9a]

$$\begin{aligned} W(\boldsymbol{x}_l \to \boldsymbol{x}_{l'}) &= \frac{1}{2\tau_{\mathrm{s}}}\left[1 - \tanh\left(\frac{\delta\mathcal{H}}{2k_{\mathrm{B}}T}\right)\right] \\ &= \frac{1}{\tau_{\mathrm{s}}} \frac{\exp\left(-\delta\mathcal{H}/k_{\mathrm{B}}T\right)}{\left[1 + \exp\left(-\delta\mathcal{H}/k_{\mathrm{B}}T\right)\right]}, \end{aligned} \tag{2.39a}$$

or

$$W(\boldsymbol{x}_l \to \boldsymbol{x}_{l'}) = \begin{cases} \dfrac{1}{\tau_{\mathrm{s}}} \exp\left(-\delta\mathcal{H}/k_{\mathrm{B}}T\right) & \text{if} \quad \delta\mathcal{H} > 0\,, \\ \dfrac{1}{\tau_{\mathrm{s}}} & \text{otherwise,} \end{cases} \tag{2.39b}$$

τ_{s} being an arbitrary factor which for the moment may be chosen as unity. (Later when we interpret the Monte Carlo process dynamically [2.9, 2.10] we shall choose τ_{s} as a unit of "Monte Carlo time" and call W a "transition probability per unit time".)

While it is easily checked that (2.39) satisfies (2.37) and (2.38), it remains to show that a sequence of states $\boldsymbol{x}_l \to \boldsymbol{x}_{l'} \to \boldsymbol{x}_{l''} \ldots$ generated with the help of (2.39) actually has the property that its probability distribution $P(\boldsymbol{x}_l)$ converges towards the canonic probability $P_{\mathrm{eq}}(\boldsymbol{x}_l)$, (2.36). A well-known plausibility argument to show this is as follows: Suppose we consider a large number of such Markov chains in parallel, and that at a given step of the process there are N_r systems in state r,

N_s systems in state s, etc., and that $\mathcal{H}(\boldsymbol{x}_r) < \mathcal{H}(\boldsymbol{x}_s)$. Using random numbers, one may construct moves $\boldsymbol{x}_r \to \boldsymbol{x}_s$, as will be discussed below. Disregarding the energy change $\delta\mathcal{H}$, the transition probability for these moves should be symmetric, i.e., $W_{\delta\mathcal{H}=0}(\boldsymbol{x}_r \to \boldsymbol{x}_s) = W_{\delta\mathcal{H}=0}(\boldsymbol{x}_s \to \boldsymbol{x}_r)$. With these "a priori transition probabilities" $W_{\delta\mathcal{H}=0}$, it is easy to construct transition probabilities which are in accord with (2.37) and (2.38), namely

$$\begin{aligned} W(\boldsymbol{x}_r \to \boldsymbol{x}_s) &= W_{\delta\mathcal{H}=0}(\boldsymbol{x}_r \to \boldsymbol{x}_s) \exp(-\delta\mathcal{H}/k_\mathrm{B}T) \\ &= W_{\delta\mathcal{H}=0}(\boldsymbol{x}_r \to \boldsymbol{x}_s) \exp\{-[\mathcal{H}(\boldsymbol{x}_s) - \mathcal{H}(\boldsymbol{x}_r)]/k_\mathrm{B}T\}, \qquad (2.40a) \\ W(\boldsymbol{x}_s \to \boldsymbol{x}_r) &= W_{\delta\mathcal{H}=0}(\boldsymbol{x}_s \to \boldsymbol{x}_r) = W_{\delta\mathcal{H}=0}(\boldsymbol{x}_r \to \boldsymbol{x}_s). \qquad (2.40b) \end{aligned}$$

The total number $N_{r\to s}$ of transitions from $\boldsymbol{x}_r$ to $\boldsymbol{x}_s$ at this step of the Markov chains is

$$\begin{aligned} N_{r\to s} &= N_r W(\boldsymbol{x}_r \to \boldsymbol{x}_s) \\ &= N_r W_{\delta\mathcal{H}=0}(\boldsymbol{x}_r \to \boldsymbol{x}_s) \exp\{-[\mathcal{H}(\boldsymbol{x}_s) - \mathcal{H}(\boldsymbol{x}_r)]/k_\mathrm{B}T\}, \qquad (2.41a) \end{aligned}$$

while the total number of inverse transitions is

$$N_{s\to r} = N_s W(\boldsymbol{x}_s \to \boldsymbol{x}_r) = N_s W_{\delta\mathcal{H}=0}(\boldsymbol{x}_r \to \boldsymbol{x}_s). \qquad (2.41b)$$

Now the net number of transitions $\Delta N_{r\to s}$ becomes

$$\begin{aligned} \Delta N_{r\to s} &= N_{r\to s} - N_{s\to r} \\ &= N_r W_{\delta\mathcal{H}=0}(\boldsymbol{x}_r \to \boldsymbol{x}_s) \left(\frac{\exp[-\mathcal{H}(\boldsymbol{x}_s)/k_\mathrm{B}T]}{\exp[-\mathcal{H}(\boldsymbol{x}_r)/k_\mathrm{B}T]} - \frac{N_s}{N_r} \right). \qquad (2.42) \end{aligned}$$

Equation (2.42) is the key result of this argument which shows that the Markov process has the desired property that states occur with probability proportional to the canonic probability $P_\mathrm{eq}(\boldsymbol{x}_l)$ as given in (2.36): As long as N_s/N_r is smaller than the ratio of the canonic probabilities we have $\Delta N_{r\to s} > 0$, i.e., the ratio N_s/N_r increases towards the ratio of canonic probabilities; conversely, if N_s/N_r is larger than the "canonic ratio", $\Delta N_{r\to s} < 0$ and hence again N_s/N_r decreases towards the correct canonic ratio. Thus asymptotically for $l \to \infty$ a steady-state distribution is reached, where N_s/N_r has precisely the value required by the canonic distribution (2.36). Instead of considering many Markov chains in parallel, we may equivalently cut one very long Markov chain into (equally long) pieces and apply the same argument to the subsequent pieces of the chain.

Now we discuss the question: What does the move $\boldsymbol{x}_l \to \boldsymbol{x}_{l'}$ mean in practice? In principle, there is enormous freedom in the choice of this move which is only restricted by the condition that the "a priori probability" $W_{\delta\mathcal{H}=0}(\boldsymbol{x}_l \to \boldsymbol{x}_{l'})$ is symmetric, $W_{\delta\mathcal{H}=0}(\boldsymbol{x}_l \to \boldsymbol{x}_{l'}) = W_{\delta\mathcal{H}=0}(\boldsymbol{x}_{l'} \to \boldsymbol{x}_l)$, and that the resulting transition probability $\tau_s W(\boldsymbol{x}_l \to \boldsymbol{x}_{l'})$ in the presence of the energy change $\delta\mathcal{H}$ should yield values

significantly different from zero and one sufficiently often. Therefore one mostly performs moves where only one (or a few) degree(s) of freedom is (are) changed, since if we change $N' \gg 1$ degrees of freedom simultaneously, we expect $\delta\mathcal{H}/k_{\mathrm{B}}T$ in (2.39) to be of the order of $N'(\varepsilon/k_{\mathrm{B}}T)$, where ε sets the energy scale [e.g., $\varepsilon = J$ for the magnetic Hamiltonians (2.1)–(2.3), and N' is equal to some power of N, i.e., $N' = N^{\delta}$. Since the temperatures of interest are such that $\varepsilon/k_{\mathrm{B}}T$ is of order unity, for nearly every move with $N' \gg 1$ we would have an extremely small transition probability if it costs energy, and hence most of the attempted moves would not be executed at all; the system "sticks" with its respective previous configuration. This clearly leads to an impractical algorithm in most cases. However, a route to overcome this problem are combined Monte-Carlo–Langevin algorithms [2.9a].

2.1.7 More About Models and Algorithms

Figure 2.4 now shows some of the moves commonly used for a variety of models under study in statistical mechanics. For the Ising model the most commonly used algorithms are the single spin-flip algorithm and the spin-exchange algorithm (Fig. 2.4a, b). Note that the single spin-flip algorithm obviously does not leave the total magnetization of the system invariant, while the spin-exchange algorithm does. Thus, these algorithms correspond to realizations of different thermodynamic ensembles: Fig. 2.4a realizes a "grand-canonical" ensemble, where temperature T and field H are the independently given thermodynamic quantities, and conjugate thermodynamic quantities (for example, the magnetization $\langle M\rangle_T$ is the quantity conjugate to the field H) need to be calculated, while Fig. 2.4b realizes a "canonical" ensemble where temperature T and magnetization M are the independently given thermodynamic quantities (and now the magnetic field $\langle H\rangle_T$ is the conjugate dependent variable which we may wish to calculate).

In calling the (T, H)-ensemble of the Ising model "grand canonical" and the (T, M)-ensemble "canonical" we apply a language appropriate to the lattice gas interpretation of the Ising model, where the spin variable S_i is reinterpreted as a local density $\varrho_i (= 0\,, 1)$ with $\varrho_i = (1 - S_i)/2$. Then $\langle M\rangle_T$ is related to the average density $\langle\varrho_i\rangle_T$ as $\langle M\rangle_T = 1 - 2\langle\varrho_i\rangle_T$, and H is related to the chemical potential of the particles which may occupy the lattice sites.

It is known that in the thermodynamic limit $N \to \infty$, different ensembles in statistical mechanics yield equivalent results. Thus the choice of the statistical ensemble and hence the associate algorithm may seem a matter of convenience. However, finite-size effects are quite different in the various ensembles, and also "rates" at which equilibrium configurations are approached in a simulation may be quite different, and hence the choice of the appropriate statistical ensemble is sometimes a delicate matter. When we use the word "rate" we have in mind already the dynamic interpretation [2.9] of the Monte Carlo process: then Fig. 2.4a realizes the Glauber [2.11] kinetic Ising model, which is a purely relaxational model without any conservation

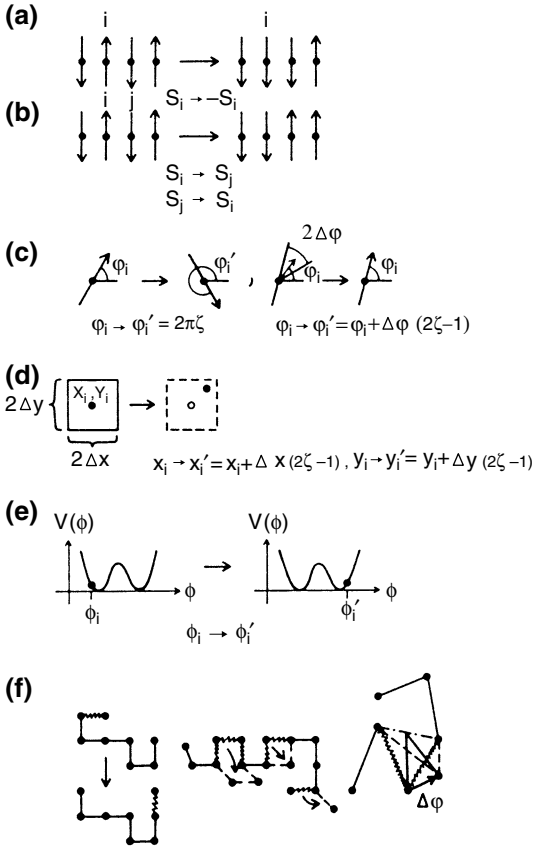


Fig. 2.4 Examples of moves $\boldsymbol{x}_l \to \boldsymbol{x}_l'$, commonly used in Monte Carlo simulations for some standard models of statistical mechanics. **a** Single spin-flip Ising model (interpreted dynamically, this is the Glauber kinetic Ising model), **b** Nearest-neighbor exchange Ising model (interpreted dynamically, this is the Kawasaki kinetic Ising model). **c** Two variants of algorithms for the *XY*-model, using a random number η equally distributed between zero and one: *left*, the angle φ_i' characterizing the new direction of the spin is chosen completely at random; *right*, φ_i' is drawn from the interval $[\varphi_i - \Delta\varphi, \varphi_i + \Delta\varphi]$ around the previous direction φ_i. **d** Moves of the coordinates of an atom in a two-dimensional fluid from its old position (x_i, y_i) to a new position equally distributed in the square of size $(2\Delta x)(2\Delta y)$ surrounding the old position. **e** Moves of a particle moving in a given single-site potential $V(\phi)$ from an old position ϕ_i to a new position ϕ_i'. **f** Moves used in the simulation of lattice models and off-lattice models of polymers (bonds which are moved are denoted by wavy lines): *left*, "slithering snake" (reptation) algorithm on the square lattice; *middle*, a dynamic algorithm relating to the Rouse model of polymer dynamics; *right*, an off-lattice algorithm for the freely jointed chain (consisting of rigid links of the same length), where two adjacent links are rotated together by a randomly chosen angle from an interval $[-\Delta\varphi, +\Delta\varphi]$ in the plane normal to the *dash-dotted axis*

laws, while Fig. 2.4b realizes Kawasaki's [2.12] kinetic Ising model which conserves magnetization.

For models with continuous degrees of freedom, such as XY (2.2) or Heisenberg magnets (2.3), but also for models of fluids (Fig. 2.4c, d), it is often advisable to choose the new degree(s) of freedom of a particle (e.g., the angle φ_i' in Fig. 2.4c, or positions x_i', y_i' in Fig. 2.4d) not completely at random, but rather in an interval around their previous values. The magnitude of this interval can then be adjusted such that the average acceptance rate for the trial moves considered in Fig. 2.4 does not become too small.

It may also be that it is inconvenient (or impossible) to sample the available phase space for a single degree of freedom uniformly. For example, while there is no difficulty in sampling angles from the interval $[0, 2\pi]$ in the left part of Fig. 2.4c, we cannot sample a variable ϕ_i from the interval $[-\infty, +\infty]$ uniformly, see Fig. 2.4e. Such a problem arises for the simulation of the so-called ϕ^4-model on a lattice:

$$\mathcal{H}_{\phi^4} = +\sum_i \left(\tfrac{1}{2}A\phi_i^2 + \tfrac{1}{4}B\phi_i^4\right) + \sum_{\langle i,j\rangle} \tfrac{1}{2}C\left(\phi_i - \phi_j\right)^2, \qquad (2.43)$$

$$-\infty < \phi_i < +\infty,$$

where A, B, C are constants, and the single-site potential $V(\phi) = \frac{1}{2}A\phi_i^2 + \frac{1}{4}B\phi_i^4$ has the familiar double-minimum shape for $A < 0$, $B > 0$. A trivial way of handling that model would be to cut off the allowed interval for ϕ_i at values which exceed the minimum position ($\phi_i^{\min} = \pm\sqrt{-A/B}$) substantially. If the potential is very steep, however, this method is rather inefficient: most of the time one would attempt to choose trial configurations ϕ_i' which are rejected because the transition probability is too small. This problem is avoided if the ϕ_i's themselves are chosen according to an importance sampling scheme already, i.e., one constructs an algorithm [2.13] which generates ϕ_i's proportional to a probability distribution

$$p(\phi_i) \propto \exp\left(-\frac{V(\phi_i)}{k_B T}\right).$$

Finally, Fig. 2.4f shows that also for random walk problems there are importance sampling alternatives to the simple-sampling techniques discussed so far. Suppose one SAW configuration of a long chain has been constructed by simple sampling. Then further configurations are generated by various "dynamic" algorithms. For example, in the "slithering snake" (reptation) [2.14] algorithm one end bond is removed from one end of the chain (it is decided at random which end is chosen) and therefore a bond is added in a random direction at the other end. Of course, this trial move is executed only if it does not violate the SAW constraints.

An alternative to this algorithm which is more realistic if one wishes to simulate polymer dynamics (e.g., the Rouse model describing the relaxation of a chain in a heat bath [2.15]) rather allows local rearrangements of groups of neighboring bonds along the chain, which may randomly flip over to new positions on the lattice, again

obeying the excluded volume restrictions (Fig. 2.4f, middle) [2.4, 2.16]. Also, the end bonds may rotate to new positions. Finally, attention is drawn to the fact that it is always possible to invent continuum (off-lattice) analogs of the models, e.g., Fig. 2.4f, right [2.17]. Again, the details of the algorithm will depend on the goals one addresses with the simulation. For example, if one wishes to study the dynamics of polymer melts [2.17] it is important to take into account the "entanglement restriction", i.e., during the random motion of the links of a chain the chain must not intersect itself or any links of other chains in the surroundings. The transition probability for motions where links would be cut then is put equal to zero, and such attempted moves are never carried out. On the other hand, if we are mainly interested in static equilibrium properties of the model, it is advantageous to define the rules of the Monte Carlo game such that the approach to equilibrium is as fast as possible. For the present problem this means one should disregard entanglement restrictions and allow the intersection of chains.

Obviously, it is impossible to exhaustively enumerate all the various possibilities that the step $\boldsymbol{x}_l \rightarrow \boldsymbol{x}_{l'}$ may mean, and how the transition probability is defined in detail. The great variability and flexibility of the Monte Carlo method allows it to be applied to many problems of very different kinds and this is clearly a major strength of this method.

2.2 Organization of Monte Carlo Programs, and the Dynamic Interpretation of Monte Carlo Sampling

2.2.1 First Comments on the Simulation of the Ising Model

Suppose now we wish to realize, as a simple example, the single spin-flip Ising model simulation of Fig. 2.4a. How is this done?

We first have to specify the type and size of the lattice and the boundary conditions which have to be used. Suppose we take a simple cubic lattice of size $L \times L \times L$ (i.e., all linear dimensions equal) and periodic boundary conditions. Then we have to specify an initial spin configuration, e.g., all spins are initially pointing up. Now one repeats again and again the following six steps:

1. Select one lattice site i at which the spin S_i is considered for flipping ($S_i \rightarrow -S_i$).
2. Compute the energy change $\delta\mathcal{H}$ associated with that flip.
3. Calculate the transition probability $\tau_s W$ for that flip.
4. Draw a random number $\mathcal{Z}$ uniformly distributed between zero and unity.
5. If $\mathcal{Z} < \tau_s W$ flip the spin, otherwise do not flip it. In any case, the configuration of the spins obtained in this way at the end of step (5) is counted as a "new configuration".
6. Analyze the resulting configuration as desired, store its properties to calculate the necessary averages. For example, if we are just interested in the (unnormalized) magnetization M_{tot}, we may update it replacing M_{tot} by $M_{\text{tot}} + 2S_i$.

More details on how one realizes an algorithm for this procedure are given in Sect. 3.4.1. It should be clear from this list that it is fairly straightforward to generalize this kind of algorithm to systems other than Ising models, such as the other possibilities considered in Fig. 2.4. The words "spin" and "spin flip(ping)" simply have to be replaced by the appropriate words for that system.

We add some comments on the practical implementation of this algorithm:

(i) By steps (3–5) the spin flip *on the average* is executed with the probability $\tau_s W$; if $\tau_s W > 1$ it is always executed. This occurs, for example, for the choice (2.39b) if $\delta\mathcal{H} < 0$. In this case the steps (3,4) need not be carried out.

(ii) Considerable freedom exists in the order in which subsequent lattice sites $\{i\}$ are selected when one repeats this process. One may go through the lattice in a regular (typewriter) fashion or one may select the lattice sites at random; for equilibrium properties this does not matter. While the random sequence of visited sites is more realistic if one wishes to obtain dynamic properties (in the spirit of the Glauber kinetic Ising model), it is somewhat slower than the regular procedure, and also needs fairly good (uncorrelated) pseudorandom numbers (with bad numbers it may happen that certain sites are never visited, and then the program output is nonsense!). It is also possible to first divide the lattice into several interpenetrating sublattices, and then go in a regular typewriter fashion first through the spins of sublattice 1, then to sublattice 2, etc. This "checker-board algorithm" [2.18] has the merit that it is straightforwardly "vectorizable" and hence can be performed very efficiently on vector computers.

(iii) Since subsequent states differ only by a single spin flip, their physical properties are very strongly correlated. If step (6), the analysis of configurations, involves many operations it may be better not to perform it after every (attempted) flip, but only after much larger "time" intervals. We define one Monte Carlo step (MCS) per site by carrying out the above (5 or 6) steps once for every lattice site, if the spins are selected consecutively. If we select them at random, the MCS/site is defined by requiring that on the average each spin is selected once. It is then advisable to perform step 6 (or time-consuming parts thereof) only once every MCS/site or even only once every τth MCS/site, where τ is some typical correlation time (which will be considered below). Also, it must be noted that although the distribution of generated states $P(\boldsymbol{x}_l)$ asymptotically (that is, for $M \to \infty$) tends to the canonic distribution $P_{\text{eq}}(\boldsymbol{x})$, (2.36) there is no reason to assume that also the states immediately following the (arbitrary!) initial configuration already have a distribution close to $P_{\text{eq}}(\boldsymbol{x})$. On the contrary, it will be necessary to generate a large number of configurations $\{\boldsymbol{x}_l\}$ until the "memory" of the initial state is lost. For example, if we start our system with its saturation magnetization but wish to simulate a temperature just above the Curie temperature, where the magnetization in equilibrium in zero magnetic field is zero, it may take a very long "time" until the initial magnetization has actually "decayed". An example of this decay towards equilibrium, as well as fluctuations in equilibrium, is given in Fig. 2.5. Again it is helpful to discuss this phenomenon using the dynamic interpretation of the Monte Carlo process, as will be shown below. In any case, it is useful to exclude from the final analysis states at the beginning of the

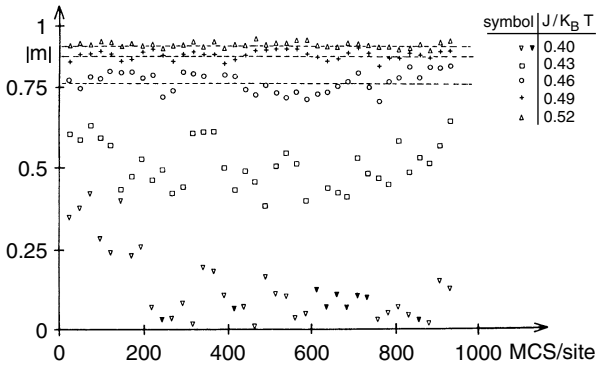


Fig. 2.5 Absolute value of the magnetization of the two-dimensional nearest neighbor Ising square lattice with ferromagnetic interaction, for lattice size $L = 55$ and periodic boundary conditions. Instantaneous values of $|m|$ are shown as a function of observation time t (in units of MCS/site) for five different temperatures. Negative values of m at $J/k_B T = 0.40$ are indicated by *filled symbols*. This shows that even in the paramagnetic region ($J/k_B T_c \approx 0.4409$) the memory of the initial starting configuration is lost only rather gradually. *Dashed horizontal straight lines* indicate the exact values of the spontaneous magnetization, from the exact solution of Yang. Note that on approaching T_c from below both the amplitude of the fluctuation δm and the relaxation time get larger (note the time needed for δm to change sign). Slightly above $T_c(J/k_B T = 0.43)$, the total running time shown here was not enough to relax the sign of the magnetization, which fluctuates around zero in a finite system at all nonzero temperatures

simulation run which are not well enough "equilibrated" (unless it is the approach towards equilibrium that one wishes to study!).

(iv) One can save computer time by storing at the beginning of the calculation the small number of different values that the transition probability W for spin flips (or spin exchange, respectively) can have, rather than evaluating the exponential function again and again. This table method works for all problems with discrete degrees of freedom, not only for the Ising model.

At very low temperatures in the Ising model, nearly every attempt to flip a spin is bound to fail. One can construct a more complicated but quicker algorithm by keeping track of the number of spins with a given transition probability W_k at each instant of the simulation. Choosing now a spin from the kth class with a probability proportional to W_k, one can make every attempted spin flip successful [2.19]. An extension of this algorithm to the spin-exchange model has also been given [2.20].

2.2.2 *Boundary Conditions*

The disturbance from the boundaries of the system is usually diminished by employing periodic boundary conditions. Thus the uppermost and the lowermost planes in three-dimensional lattice are regarded as neighbors in the energy calculation, as are

the back and front planes, and also the leftmost and rightmost planes of the lattice. This is shown in Fig. 2.6a for the example of 6×6 square lattice. In the programming example reproduced in Table 2.1, periodic boundary conditions are explicitly implemented. Note, however, that the choice of appropriate linear dimensions and boundary conditions has to be done with a little bit of thought. For example, while for the ferromagnetic Ising model the linear dimension L may be either odd or even, for the antiferromagnetic Ising model it must be even, because otherwise the two-sublattice structure of the cubic Ising antiferromagnet would not fit to the lattice. Clearly, the periodic boundary condition must have disturbing effects in cases where the system wants to develop a long-range order that is not commensurate with the linear dimension of the box, which necessarily occurs in models exhibiting commensurate–incommensurate phase transitions, such as the ANNNI (anisotropic next-nearest neighbor Ising) model [2.22]. In addition, at all second-order phase transitions the critical divergence of the correlation length is strongly disturbed by the finite size and the periodicity of the system. These finite-size and boundary effects on phase transitions have received attention over a long period of time [2.23–2.27] and will be treated in Sect. 2.3. A variant of the periodic boundary condition is often applied if one stores the labels of the lattice sites in a one-dimensional array going in a typewriter fashion through the (simple cubic) lattice. Then the nearest neighbors of site i are taken as $i \pm 1, i \pm L, i \pm L^2$, which implies a skewed periodic boundary condition. This case is also illustrated in Fig. 2.6a. However, it is possible to take advantage of the one-dimensional labeling (useful particularly in the context of a good "vectorization" of Monte Carlo programs to be run on vector processors) and retain the strictly periodic boundary condition with some programming effort [2.28].

Sometimes boundary conditions other than fully periodic boundary conditions are used, due to the intention to study properties other than bulk properties of the system. For example, in order to study free surfaces of a system the so-called "free boundary conditions" are used. Then the "missing spins" next to the free surface are simply set equal zero. This case is illustrated in Fig. 2.6b. It may then be advisable to spend more statistical effort in the surface layer (and its adjacent layers) rather than in layers deep in the bulk. Such an algorithm is called preferential surface site selection [2.29]. In order to simulate a small superparamagnetic particle one may use free surfaces in all lattice directions [2.30–2.32] and also rather arbitrary shapes of these systems can be prescribed [2.33]. In addition, one may wish to apply boundary fields, particularly if simulations of wetting and layering phenomena are desired [2.34–2.36]. If one wishes to simulate a thin-film geometry, one uses free boundary conditions in one lattice direction and periodic boundary conditions in the remaining ones, see Fig. 2.6b [2.29–2.37]; a semi-infinite geometry is simulated by choosing one free surface and a self-consistent effective field [2.21] on the opposite surface, while in the remaining directions again periodic boundary conditions are used [2.38], see Fig. 2.6c.

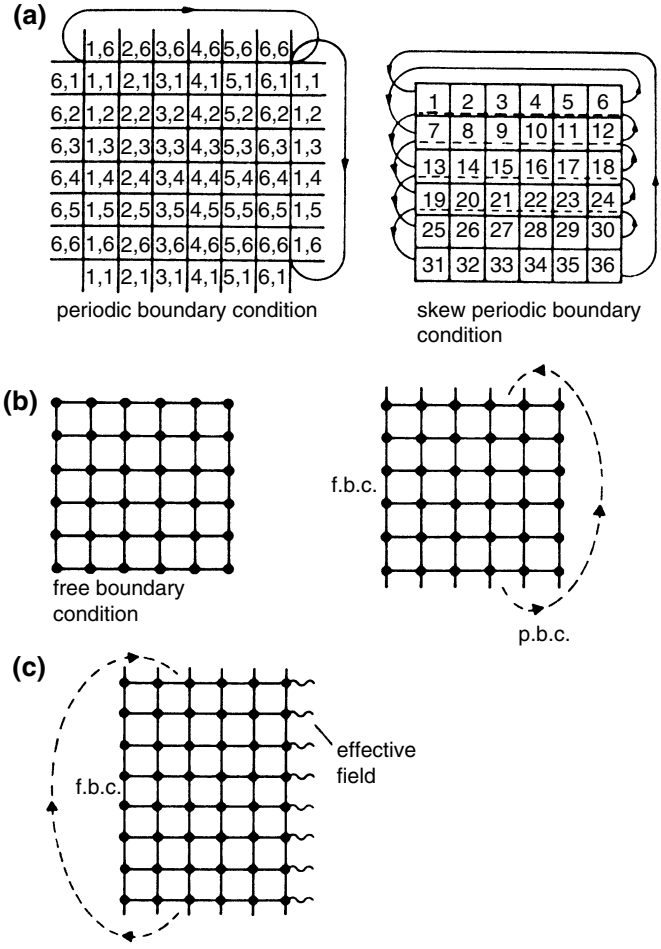


Fig. 2.6 **a** Square 6×6 lattice with periodic boundary conditions (*left*) and with skew periodic boundary conditions (*right*). For the periodic boundary conditions, each lattice site i is labeled by its two Cartesian coordinates x_i, y_i, which are integers when the lattice spacing is taken to be unity. For the skew periodic boundary condition, the sites are labeled in the typewriter fashion as indicated. **b** Square 6×6 lattice (lattice sites being denoted by dots, nonvanishing bonds by straight lines connecting them) with free boundary conditions in all lattice directions (*left*) or with free boundary conditions in one lattice direction and periodic boundary conditions in the other lattice direction (*right*). This latter choice simulates a thin-film geometry. **c** Boundary conditions simulating a semi-infinite system. A 6×8 square lattice has one boundary (eight spins) free, the opposite one is exposed to an effective boundary field. The magnitude of this field is adjusted such that the magnetization in this boundary layer takes on its bulk value [2.21]. In the other lattice direction, periodic boundary conditions are used

Table 2.1 Example of a program for the Ising model with a simple cubic lattice. `srand()` initializes the random number generator `rand()`. The energy change is calculated as discussed in Sect. 2.2.1, the flipping probability is taken as $\exp(-\Delta E/k_B T)$

```
#include <stdio.h>
#include <stdlib.h>
#include <math.h>

# define L  100

int main(int argc, char *argv[]){

    int spin[L][L][L];
    int i,j,k;
    int ip1,im1,jp1,jm1,kp1,km1;
    int mcs;
    int spinValue, spinNeighbourSum;
    int mag;
    int mcsmax;
    int seed;

    double r;
    double K;
    double energyChange;
    double rm;
    double averageMag;
    double TTC;

    /***** Set the parameters *****/
    mcsmax = 1000;
    TTC    = 1.2;
    K      =  0.2216544 / TTC;
    seed   = 4711;

    srand(seed);

    /***** Initialize the lattice *****/
    for (i=0;i<L;i++) {
        for (j=0;j<L;j++) {
            for (k=0;k<L;k++) {
                spin[i][j][k] = -1;
            }
        }
    }
    mag        = - L*L*L;
    averageMag = 0.0;

    /********** Monte Carlo Loop ***********/
    for (mcs=0;mcs<mcsmax;mcs++) {

        /***** Loop over all lattice sites *****/
        for (i=0;i<L;i++) {
            for (j=0;j<L;j++) {
                for (k=0;k<L;k++) {

                    /***** Periodic boundary conditions *****/
                    ip1 = i+1 >= L ? i+1-L : i+1;
                    im1 = i-1 <  0 ? i-1+L : i-1;
                    jp1 = j+1 >= L ? j+1-L : j+1;
                    jm1 = j-1 <  0 ? j-1+L : j-1;
                    kp1 = k+1 >= L ? k+1-L : k+1;
                    km1 = k-1 <  0 ? k-1+L : k-1;

                    /***** Change in energy *****/
                    spinValue        = spin[i][j][k];
                    spinNeighbourSum = spin[im1][j][k] + spin[ip1][j][k] +
                                       spin[i][jm1][k] + spin[i][jp1][k] +
                                       spin[i][j][km1] + spin[i][j][kp1];
```

(continued)

Table 2.1 (continued)

```
                    energyChange     = 2.0 * K * spinValue * spinNeighbourSum;

                    /***** Generate random number *****/
                    r = (float)rand() / (float)RAND_MAX;

                    /***** Accept/reject the change *****/
                    if (energyChange==0.0 || r<exp(-(double) energyChange)) {
                        spin[i][j][k] = -spinValue;
                        mag           -= 2 * spinValue;
                    }
                }
            }
        }
        rm =  (double) mag / (float) (L*L*L);
        averageMag += fabs(rm);
    }

    averageMag /= (float) mcsmax;
    printf("%f\n",averageMag);

    return 0;
}
```

2.2.3 *The Dynamic Interpretation of the Importance Sampling Monte Carlo Method*

We now turn to a discussion of the correlations between the configurations generated sequentially one after the other in the Markov chain. Clearly, these correlations strongly affect the accuracy that can be obtained with a given number of total steps by the Monte Carlo program. These correlations can be understood by a *dynamic interpretation* of the Monte Carlo averaging in terms of a master equation describing a well-defined dynamic model with stochastic kinetics [2.8–2.10, 2.39]. Not only is the interpretation of correlations as time correlations useful to the understanding of accuracy, it is also the theoretical basis for the application of Monte Carlo methods to the simulation of dynamic processes [2.40–2.42]. These dynamic applications include such diverse fields as the Brownian motion of macromolecules [2.4, 2.43], relaxation phenomena in spin glasses [2.41] and quadrupolar glasses [2.44], nucleation phenomena [2.40–2.45] and spinodal decomposition of mixtures [2.40, 2.45], diffusion-limited aggregation and related irreversible growth phenomena [2.42, 2.46], and diffusion in alloys and at surfaces [2.40, 2.47]. At this point, we just associate a "time" t with the scale ν of the subsequent configurations. We may normalize the time scale such that $N\tau_s^{-1}$ single-particle transitions are performed in unit time [in (2.39) we have already introduced this factor τ_s^{-1} into the transition probability]. Then the time unit is 1 MCS (Monte Carlo step per particle). We consider the probability $P(\boldsymbol{x}_\nu) \equiv P(\boldsymbol{x}, t)$ that at time t a configuration $\boldsymbol{x}$ occurs in the Monte Carlo process. This probability satisfies the Markovian master equation [2.8–2.10, 2.48]

$$\frac{\mathrm{d}P(\boldsymbol{x},t)}{\mathrm{d}t} = -\sum_{\boldsymbol{x}'} W(\boldsymbol{x}\to\boldsymbol{x}')P(\boldsymbol{x},t) + \sum_{\boldsymbol{x}'} W(\boldsymbol{x}'\to\boldsymbol{x})P(\boldsymbol{x}',t). \qquad (2.44)$$

Equation (2.44) describes the balance considered already above (2.40)–(2.42) by a rate equation, the first sum on the right hand side representing all processes where one moves away from the considered state $\boldsymbol{x}$ (and hence its probability is decreased), while the second sum contains all reverse processes (which hence lead to an increase of the probability of finding $\boldsymbol{x}$). In thermal equilibrium the detailed balance condition (2.37) ensures that these two sums always cancel, and hence for $P(\boldsymbol{x},t) = P_{\mathrm{eq}}(\boldsymbol{x})$ we have $dP(\boldsymbol{x},t)/dt = 0$, as is required. In fact, $P_{\mathrm{eq}}(\boldsymbol{x})$ is the steady-state distribution of the above master equation. If the potential energy is finite for arbitrary configurations $\{\boldsymbol{x}\}$ of the system, we can conclude from the finiteness of the system that it must be ergodic. However, as soon as we have infinite potentials (such as in the self-avoiding walk problem, Figs. 2.1c and 2.4f), certain configurations $\boldsymbol{x}$ are forbidden, and then ergodicity may be a problem. Even in finite systems the phase space may decompose into several valleys or "pockets" which are mutually inaccessible. There is no general rule about whether this happens or not, it really depends on the details of the algorithm. For example, in the case of the dynamic simulations of SAWs on lattices (Fig. 2.4f) it can be shown (see, e.g., [2.49]) that certain configurations (see Fig. 2.7 for an example) are inaccessible with the algorithms shown there, although in practice this is not a problem at all since the statistical weight of these inaccessible states is negligibly small for the averages of interest [2.50]. But this point certainly warrants some care. In practice one may find an apparent "breaking of ergodicity" even for systems which are ergodic, if the "time" over which the averaging is extended is not long enough, i.e., less than some so-called "ergodic time" τ_{e} [2.51]. This ergodicity breaking is intimately related to spontaneous symmetry breaking associated with phase transitions in the system. In a strict sense, these phase transitions can occur only in the thermodynamic limit $N\to\infty$, and hence τ_{e} also diverges as $N\to\infty$. The finiteness of τ_{e} in the regime where spontaneous order occurs for $N\to\infty$ is also a finite-size effect. We shall return to this problem in our next section where finite-size effects of various kinds will be discussed, and assume for the moment that $\lim P(\boldsymbol{x},t) = P_{\mathrm{eq}}(\boldsymbol{x})$, i.e., the ergodicity property, can be verified on practically accessible time scales.

In (2.44) we have written $\mathrm{d}P(\boldsymbol{x},t)/\mathrm{d}t$ rather than $\Delta P(\boldsymbol{x},t)/\Delta t$, i.e., we work with differentials rather than discrete differences. This point is sometimes criticized, e.g., in [2.52], where it is suggested that due to the discrete time increment $\Delta t = \tau_{\mathrm{s}}/N$ one should rather view the Monte Carlo dynamics as a discrete map, a problem in "nonlinear dynamics", predicting hence the possibility of chaotic motions, limit cycles, etc. However, we feel that this criticism is not relevant. As pointed out in [2.53], the time increment Δt should not be considered as a constant, but is instead a continuous variable stochastically fluctuating with distribution $(N/\tau_{\mathrm{s}})\exp[-\Delta t N/\tau_{\mathrm{s}}]$ which has a mean value $\overline{\Delta t} = \tau_{\mathrm{s}}/N$. Since the time scale on which dynamic correlations decay is of the order of τ_{s} itself or even larger, these fluctuations of the time variable relative to the "time" proceeding in regular steps $\Delta t = \tau_{\mathrm{s}}/N$ labeling

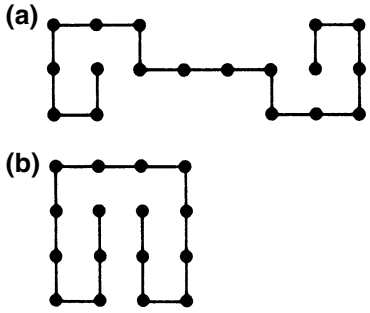


Fig. 2.7 Examples of self-avoiding walk configurations on the square lattice which are inaccessible by dynamic Monte Carlo algorithms. Case **a** is a configuration which is inaccessible by the "slithering snake" algorithm, and also cannot relax by this algorithm. Case **b** can relax neither by the slithering snake algorithm nor by the "kink-jump" method

the Monte Carlo microsteps are averaged out when one calculates time-displaced correlation functions.

Thus the average obtained in (2.35) can simply be interpreted as a time average along the stochastic trajectory in phase space, controlled by the master equation (2.44) of the system, i.e., for the variable A considered in (2.5) and (2.35) we now get

$$\overline{A} = \frac{1}{t_M - t_{M_0}} \int_{t_{M_0}}^{t_M} A(t)\,\mathrm{d}t, \tag{2.45}$$

where t_M is the "time" elapsed after M configurations have been generated, t_{M_0} the time after $M_0 < M$ configurations have been generated ($t_M = M\tau_s/N$, $t_{M_0} = M_0\tau_s/N$), and we have anticipated that the first M_0 configurations are actually omitted from the average in (2.35), which therefore actually reads

$$\overline{A} = \frac{1}{(M - M_0)} \sum_{\nu=M_0+1}^{M} A(\boldsymbol{x}_\nu).$$

Since $\boldsymbol{x}_\nu$ is the configuration $\boldsymbol{x}(t)$ appearing at the time $t = t_\nu = \nu\tau_s/N$, we can consider $A = A(\boldsymbol{x}(t))$ simply as a function of system time t itself. Comparing the *time average* (2.45) with our starting formula (2.5) which is a canonic *ensemble average*, it is obvious that this importance sampling Monte Carlo method leads to a problem of *ergodicity*, i.e., the question of whether the time average is identical to the ensemble average, as anticipated above.

It is now also quite obvious how we can define time-displaced correlation functions $\langle A(t)B(0)\rangle_T$ or $\overline{A(t)B(0)}$, where B stands symbolically for any other physical variable:

$$\overline{A(t)B(0)} = \frac{1}{t_M - t - t_{M_0}} \int_{t_{M_0}}^{t_M - t} A(t + t')B(t')\mathrm{d}t', \quad t_M - t > t_{M_0}. \tag{2.46}$$

In practice, (2.45) and (2.46) are used for times t_{M_0} large enough such that the system has relaxed towards equilibrium during the time t_{M_0}, and then the states $\boldsymbol{x}(t)$ included in the sampling from t_{M_0} to time t_M are already distributed according to the equilibrium distribution, $P(\boldsymbol{x}, t) = P_{\mathrm{eq}}(\boldsymbol{x})$, independent of time. However, it is also interesting to study the nonequilibrium relaxation process by which equilibrium is approached. In this region $A(t) - \overline{A}$ is systematically dependent on the observation time t, and an ensemble average $\langle A(t)\rangle_T - \langle A(\infty)\rangle_T [\lim_{t\to\infty} \overline{A} = \langle A\rangle_T = \langle A(\infty)\rangle_T$ if the system is ergodic] is nonzero. Hence we define

$$\langle A(t)\rangle_T = \sum_{\{\boldsymbol{x}\}} P(\boldsymbol{x}, t)A(\boldsymbol{x}) = \sum_{\{\boldsymbol{x}\}} P(\boldsymbol{x}, 0)A(\boldsymbol{x}(t)), \tag{2.47}$$

where in the second step of this equation we have used the fact that the ensemble average involved is actually an *average weighted by* $P(\boldsymbol{x}, 0)$ *over an ensemble of initial states* $\boldsymbol{x}(t = 0)$, which then evolve as described by the master equation (2.44). In practice, (2.47) means an average over $n_{\mathrm{run}} \gg 1$ statistically independent runs,

$$[A(t)]_{\mathrm{av}} = \frac{1}{n_{\mathrm{run}}} \sum_{l=1}^{n_{\mathrm{run}}} A(t, l), \tag{2.48}$$

where $A(t, l)$ is the observable A observed at time t in the lth run of this nonequilibrium Monte Carlo averaging. These runs in practice differ by use of different random numbers for each realization of the time evolution, and by use of different initial configurations $\boldsymbol{x}(t = 0)$. (In special cases, however, such as the study of the decay of the magnetization of an Ising ferromagnet heated from zero temperature to a nonzero temperature, the initial configuration is unique and hence common to all the runs.)

Before going further we now ask the question: Is the "time" t associated with the label ν of subsequent states generated by the Monte Carlo sampling related to the physical time by which a real system evolves? In general the answer is that this is *not* the case. Systems like Heisenberg magnets, classical fluids, etc., do have a time evolution described by deterministic equations for their variables. For example, Newton's laws, not the master equation (2.44), describe the motions of molecules in a fluid, and the Heisenberg ferromagnet considered in (2.3) evolves according to the equation

$$\hbar \frac{\mathrm{d}}{\mathrm{d}t} S_k^z(t) = i \left[S_k^z, \mathcal{H}_{\mathrm{Heisenberg}}\right] = -2J \sum_{j(\neq k)} \left[S_k^y S_j^x - S_k^x S_j^y\right], \tag{2.49}$$

where $2\pi\hbar$ is Planck's constant and j is a nearest-neighbor of site k. Although both the (artificial) stochastic dynamics and the actual physical dynamics lead to

the same thermal equilibrium distribution for $N \to \infty$ (and for finite N as well, if they correspond to exactly the same statistical ensemble, namely a microcanonical ensemble where the energy is strictly conserved), there is in general little similarity between the stochastic dynamics described by (2.44) and the actual dynamics. For example, for temperatures less than the Curie temperature, (2.49) leads to the well-known spin wave excitations, while (2.44) can never yield any propagating modes, but only simple relaxational behavior.

The situation for the Ising Hamiltonian (2.1) is different, of course, since $[S_k, \mathcal{H}_{\text{Ising}}] \equiv 0$: it does not provide any intrinsic time evolution. For the Ising model, the stochastic kinetics provided by (2.44) can be interpreted physically in terms of a very weak coupling of the spins to a heat bath (the thermal vibrations of an underlying crystal lattice, for instance), which induces random spin flips in the system. Similarly, in an interstitial alloy the diffusion of the interstitial atoms may be modeled by a stochastic hopping between the available lattice sites [2.40, 2.54, 2.55]. Since the mean time between two successive jumps is orders of magnitude larger than the time scale of atomic vibrations in the solid, the phonons can again be reasonably well approximated as a heat bath, as far as the diffusion is concerned. Of course, there are also cases of interstitial alloys where this approximation gets inaccurate, such as superionic conductors. For the realistic simulation of the dynamics of such systems, the molecular dynamics method [2.56] where one integrates Newton's laws numerically is an alternative. The molecular dynamics method gets into trouble, however, if the system contains two sets of degrees of freedom with very different characteristic times, because the integration time step must be much less than the characteristic time of the fast degrees of freedom, and it may be inconvenient (or even impossible) to extend the time scale of the numerical integration up to the relaxation time of the slow degrees of freedom. The latter can then often be modeled much more efficiently by a suitable Monte Carlo algorithm.

2.2.4 *Statistical Errors and Time-Displaced Relaxation Functions*

Apart from applications studying the dynamics of suitable model systems, the dynamic interpretation provided by (2.44)–(2.48) is very useful for understanding the "statistical errors" of Monte Carlo sampling [2.9]. It is on this point that we now focus.

Suppose n successive observations A_μ, $\mu = 1, \ldots, n$, of a quantity A have been stored, with $n \gg 1$. We consider the expectation value of the square of the statistical error

$$\langle(\delta A)^2\rangle = \left\langle\left[\frac{1}{n}\sum_{\mu=1}^{n}\left(A_\mu - \langle A\rangle\right)\right]^2\right\rangle = \frac{1}{n^2}\sum_{\mu=1}^{n}\langle\left(A_\mu - \langle A\rangle\right)^2\rangle + \frac{2}{n^2}\sum_{\mu_1=1}^{n}\sum_{\mu_2=\mu_1+1}^{n}\left(\langle A_{\mu_1}A_{\mu_2}\rangle - \langle A\rangle^2\right). \tag{2.50}$$

Changing the summation index μ_2 to $\mu_2 + \mu$, (2.50) can be rewritten as

$$\langle(\delta A)^2\rangle = \frac{1}{n}\left[\langle A^2\rangle - \langle A\rangle^2 + 2\sum_{\mu=1}^{n}\left(1-\frac{\mu}{n}\right)\left(\langle A_0 A_\mu\rangle - \langle A\rangle^2\right)\right]. \tag{2.51}$$

Now we remember that a time $t_\mu = \delta t\,\mu$, is associated with the Monte Carlo process, δt being the time interval between two successive observations A_μ, $A_{\mu+1}$. [It is possible to take $\delta t = \Delta t = \tau_s/N$, i.e., every Monte Carlo microstep is included in the calculation of $\langle(\delta A)^2\rangle$, but often it is more efficient to take δt much larger than Δt, e.g., $\delta t = \tau_s$ or $\delta t = 10\,\tau_s$, etc.] Transforming the summation into a time integration and dropping the index μ from t_μ, we obtain [2.9]

$$\begin{aligned}\left\langle(\delta A)^2\right\rangle &= \frac{1}{n}\left[\langle A^2\rangle - \langle A\rangle^2 + 2\frac{1}{\delta t}\int_0^{t_n}\left(1-\frac{t}{t_n}\right)\left[\langle A(0)A(t)\rangle - \langle A\rangle^2\right]\mathrm{d}t\right] \\ &= \frac{1}{n}\left(\langle A^2\rangle - \langle A\rangle^2\right)\left[1 + \frac{2}{\delta t}\int_0^{t_n}\left(1-\frac{t}{t_n}\right)\frac{\langle A(0)A(t)\rangle - \langle A\rangle^2}{\langle A^2\rangle - \langle A\rangle^2}\mathrm{d}t\right].\end{aligned} \tag{2.52}$$

Next we denote the normalized relaxation function of the quantity A as $\phi_A(t)$ and define it as

$$\phi_A(t) = \frac{\langle A(0)A(t)\rangle - \langle A\rangle^2}{\langle A^2\rangle - \langle A\rangle^2}. \tag{2.53}$$

Note that $\phi_A(0) = 1$ and $\phi_A(t)$ decays to zero as $t \to \infty$. Let us assume that $\phi_A(t)$ has actually decayed to zero essentially already on a time-scale τ_A, with $\tau_A \ll t_n$. We can give a precise meaning to τ_A in terms of the integral

$$\tau_A = \int_0^\infty \phi_A(t)\mathrm{d}t. \tag{2.54}$$

Since we have assumed that $\phi_A(t)$ differs from zero appreciably only for times $t \ll t_n$, the term t/t_n in (2.52) can be neglected in comparison with unity, and the upper limit of the integration replaced by infinity. This yields

$$\langle(\delta A)^2\rangle = \frac{1}{n}\left[\langle A^2\rangle - \langle A\rangle^2\right]\left(1 + 2\frac{\tau_A}{\delta t}\right). \tag{2.55}$$

If $\delta t \gg \tau_A$, then the parenthesis in (2.55) is unity to a very good approximation, and the statistical error has just the same form as encountered for simple sampling [cf. (2.26)]. In the inverse case, where $\delta t \ll \tau_A$, we find instead ($n\delta t = \tau_{\text{obs}}$ is the time over which the averaging is extended)

$$\langle(\delta A)^2\rangle \approx \frac{2\tau_A}{n\delta t}\left[\langle A^2\rangle - \langle A\rangle^2\right] = 2\frac{\tau_A}{\tau_{\text{obs}}}\left[\langle A^2\rangle - \langle A\rangle^2\right], \tag{2.56}$$

which shows that then the *statistical error* is *independent* of the choice of the *time interval* δt. Although for a given averaging time t_n a choice of a smaller value δt results in a correspondingly larger value of the number n of observations, it does not decrease the statistical error; only the ratio between the relaxation time τ_A and the observation time τ_{obs} matters. The fact that $\langle(\delta A)^2\rangle$ in general is not given by the simple sampling result $\left[\langle A^2\rangle - \langle A\rangle^2\right]/n$, but is enhanced by some factor, had been noticed before a proper dynamic interpretation of importance sampling was known and therefore the enhancement factor was called the "statistical inefficiency" of the method [2.57]. Obviously, this is not a useful notion since the factor $1 + 2(\tau_A/\delta t)$ necessarily results from the fact that Markov processes are described by associated master equations. Conversely, the dynamic view as exposed here leads to the idea that one may exploit the freedom in the choice of the transition probability $W(\boldsymbol{x} \to \boldsymbol{x}')$ as well as in the microscopic nature of the meaning of a transition $\boldsymbol{x} \to \boldsymbol{x}'$ to make τ_A as small as possible. This idea is a subject of much research [2.58], and will not be followed up here. If it can be successfully implemented, the usefulness of Monte Carlo methods near phase transitions (where τ_A diverges in the thermodynamic limit: critical slowing down [2.59]) would be substantially improved, see Chap. 4.

We conclude this section by defining a nonlinear relaxation function $\phi_A^{nl}(t)$ in terms of the nonequilibrium average (2.47)

$$\phi^{(nl)}(t) = \frac{\langle A(t)\rangle_T - \langle A(\infty)\rangle_T}{\langle A(0)\rangle_T - \langle A(\infty)\rangle_T} \tag{2.57}$$

and its associate time

$$\tau_A^{(nl)} = \int_0^\infty \phi_A^{(nl)}(t)\,\mathrm{d}t. \tag{2.58}$$

The condition that the system is well equilibrated then simply reads

$$t_{M_0} \gg \tau_A^{(nl)}. \tag{2.59}$$

Note that (2.59) must hold for all quantities A, and hence it is important to focus on the slowest-relaxing quantity (for which $\tau_A^{(nl)}$ is largest) if one wishes to estimate the suitable choice of t_{M_0} reliably. Near second-order phase transitions, the slowest-relaxing quantity is usually the order parameter of the transition, and not the internal energy. Hence the "rule" published in some Monte Carlo investigations that the

equilibration of the system is established by monitoring the time evolution of the internal energy is clearly not a valid procedure.

2.3 Finite-Size Effects

2.3.1 Finite-Size Effects at the Percolation Transition

The simplest phase transition problem in statistical physics is probably the purely geometric problem of the so-called percolation transition [2.60]. One considers an *infinite* lattice where each site is randomly occupied with probability p and empty with probability $1 - p$ (site percolation problem [2.60]). Neighboring occupied sites are said to form "clusters" (Fig. 2.8). There exists a critical concentration p_c such that for $p < p_c$ there exist only clusters of finite size l on the lattice $\{l = 1, 2, \dots\}$, while for $p \geq p_c$ an infinite cluster has formed that "percolates" from one boundary of the lattice to the opposite one.

For a *finite* lattice of linear dimension L (to be specific, we take it simple cubic) it is straightforward to write down an algorithm which generates sample configurations of such a partially filled lattice. If we denote the occupation variable of a lattice site (i, j, k) by $N(i, j, k)$ and the C procedure `rand()` supplies random numbers equally distributed, a Monte Carlo program for this problem is as follows:

```
for (i = 0; i < L; i++){
    for (j = 0; j < L; j++){
        N[i][j] = 0;
        r = (float) rand() / (float) RAND_MAX;
        if (r < p) {
            N[i][j] = 1;
        }
    }
}
```

One "sweep" through the lattice determines the whole system; there is no need to wait until some "equilibrium" is established, unlike for the importance-sampling method discussed in the previous section, but we again have a simple-sampling problem. Of course, to answer questions such as: How many clusters $n_l(p)$ containing l occupied sites exist in the lattice per lattice site? Does there exist a "spanning cluster" of occupied sites which reaches from one boundary to the opposite one? etc., one needs sophisticated programs to analyze the generated configurations, which are not considered here [2.61, 2.62] but will be discussed in Sect. 3.2.4, as well as various generalizations and variants of the percolation problem.

For the description of the percolation transition, we not only want to estimate p_c, but also compute quantities like the percolation probability $P_\infty(p)$ (the probability that an occupied site is part of the percolating cluster) and the percolation susceptibility χ defined as

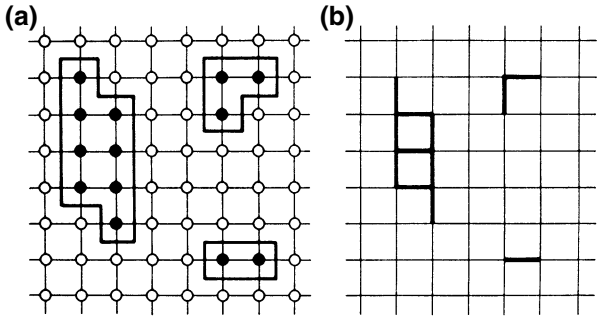


Fig. 2.8 **a** Clusters on the square lattice for the site percolation problem. Sites are occupied with probability p (*dots*) or empty with probability $1-p$ (*circles*). Neighboring occupied sites form clusters. The 8×8 lattice shown here hence contains clusters with $l=2, l=3$ and $l=8$ occupied sites. **b** Clusters on the square lattice for the bond percolation problem. Bonds are occupied with probability p (*thick lines*) or empty with probability $1-p$ (*thin lines*). Occupied bonds which are connected to each other form clusters. The lattice shown contains clusters with $l=1$, $l=2$ and $l=9$ occupied *bonds*

$$\chi = {\sum_{l=1}^{\infty}}' l^2 n_l(p)/p. \tag{2.60}$$

(The prime means that the largest cluster is omitted from the summation.) These quantities are expected to show the following critical singularities (in the *infinite* lattice) for $|p-p_c| \to 0$ [2.60]:

$$P_\infty(p) = \hat{B}_p \left(\frac{p}{p_c}-1\right)^{\beta_p}, \quad p > p_c \, [P_\infty(p) \equiv 0 \text{ for } p < p_c], \tag{2.61a}$$

$$\chi(p) = \begin{cases} \hat{\Gamma}_p^+ \left(1-\dfrac{p}{p_c}\right)^{-\gamma_p}, & p < p_c, \\ \hat{\Gamma}_p^- \left(\dfrac{p}{p_c}-1\right)^{-\gamma_p}, & p > p_c. \end{cases} \tag{2.61b}$$

In a *finite* lattice, $\chi(p)$ cannot diverge but reaches a maximum of finite height only; the magnitude of this maximum depends on the size of the lattice (Fig. 2.9) [2.63]. Similarly, the percolation probability $P_\infty(p)$ cannot vanish at any $p > 0$, but must attain small nonzero values as soon as $p > 0$. [Even the probability that all sites are occupied is nonzero in a finite system as soon as $p > 0$, namely it is $p^{L^d} = \exp(L^d \ln p)$ in a d-dimensional system; percolation occurs with the probability $p^L = \exp(L \ln p)$ as $p \to 0$]. Thus in a finite lattice the singularities [(2.61a) and (2.61b)] associated with the percolation transition are smoothed out, and this rounding at the transition is easily understandable geometrically. On a finite

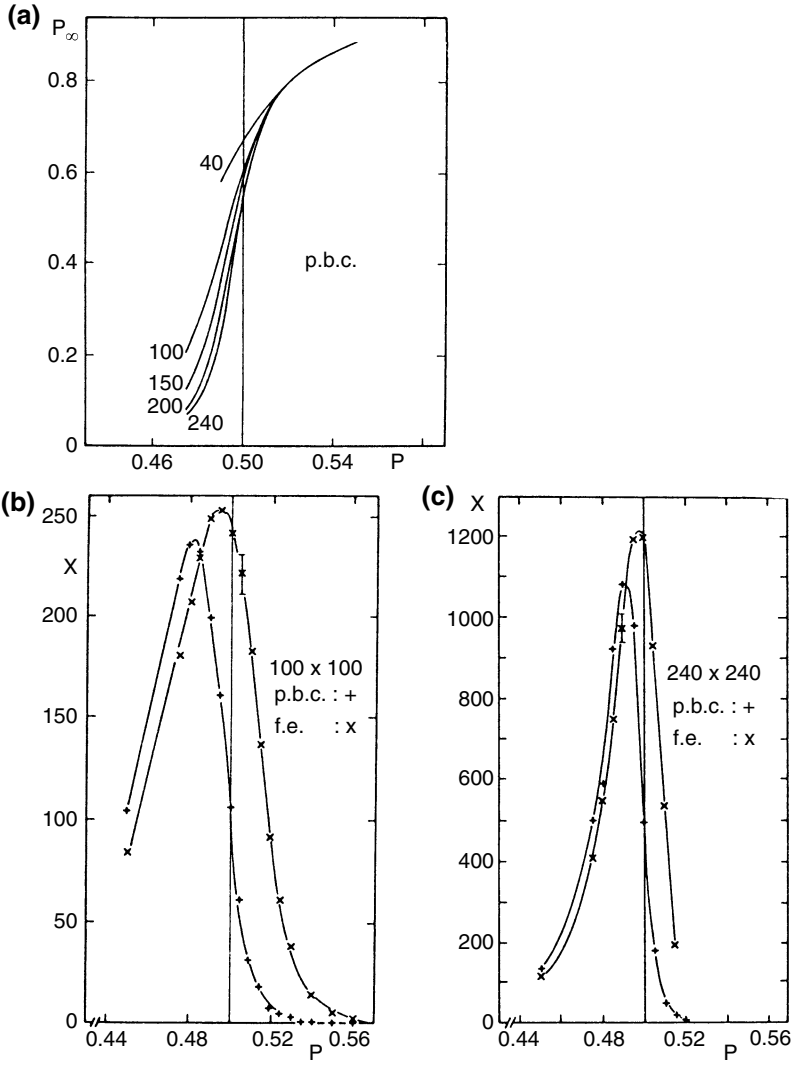


Fig. 2.9 Percolation probability $P_\infty^{(L)}(p)$ plotted versus p (**a**), for the bond percolation problem on the square lattice, using various choices of L and periodic boundary conditions, **b** and **c** show corresponding results for $\chi^{(L)}(p)$, including also data with free edges (f.e.). (From [2.63])

lattice only finite clusters are possible, and both the cluster size distribution $n_l(p)$ and the percolation probability $P_\infty(p)$ are smooth (analytic) functions of p. For the infinite lattice, the cluster size distribution behaves as [2.60]

$$n_l(p) \underset{l\to\infty}{=} l^{-\tau}\tilde{n}\left\{l^\sigma\left(1-\frac{p}{p_c}\right)\right\}, \tag{2.62a}$$
$$\tau = 2 + 1/\delta_p,$$
$$\sigma = 1/\left(\beta_p\delta_p\right) = 1/\left(\gamma_p + \beta_p\right).$$

While on the infinite lattice at p_c the cluster size distribution decays according to a power law, $n_l(p) = l^{-\tau}\tilde{n}(0)$ as $l \to \infty$, on a finite lattice this power law holds only for clusters whose radii r_l are distinctly smaller than the lattice linear dimension L. Since

$$r_l \underset{l\to\infty}{=} \hat{r}l^{1/d_f}, \tag{2.62b}$$

where d_f is the "fractal dimensionality" [2.60] of the percolation clusters, and at $p = p_c$ the probability that a cluster percolates ($r_l \approx L$) is unity, we conclude

$$L^d\int_{r_l=L}^{\infty} n_l(p)\mathrm{d}l \approx L^d\tilde{n}(0)\int_{(L/\hat{r})^{d_f}}^{\infty} l^{-\tau}\mathrm{d}l$$
$$= \tilde{n}(0)L^{d+d_f(1-\tau)}\hat{r}^{-d_f(1-\tau)}/(\tau-1) = 1, \tag{2.62c}$$

which implies that the exponent of L vanishes, i.e., $d_f = d/(\tau-1) = d/(1+1/\delta_p) = d\beta_p\delta_p/(\beta_p\delta_p+\beta_p) = d(\beta_p+\gamma_p)/(2\beta_p+\gamma_p) = (\beta_p+\gamma_p)/\nu_p = d - \beta_p/\nu_p$, where scaling laws such as the hyperscaling relation involving the correlation length exponent $\nu_p(\xi_p \propto |p-p_c|^{-\nu_p}), d\nu_p = 2\beta_p + \gamma_p$, is used. From this expression for d_f and (2.62b) we conclude that the number of sites in a "spanning cluster" with $r_l = L$ is $l_L = (L/\hat{r})^{d_f} \propto L^{d-\beta_p/\nu_p}$, and hence the fraction of occupied sites belonging to such a spanning cluster must be of the order of $P_\infty^{(L)}(p_c) = L^{-d}l_L \propto L^{-\beta_p/\nu_p}$.

2.3.2 *Finite-Size Scaling for the Percolation Problem*

Now an expression which interpolates between (2.61a) for $L \gg \xi_p$ and the above result for $P_\infty^{(L)}(p)$ at $p = p_c$ is obtained by the finite-size scaling hypothesis [2.23–2.26] [note that $P_\infty^{(L)}(p)$ may also be defined as the fraction of sites in the largest cluster]

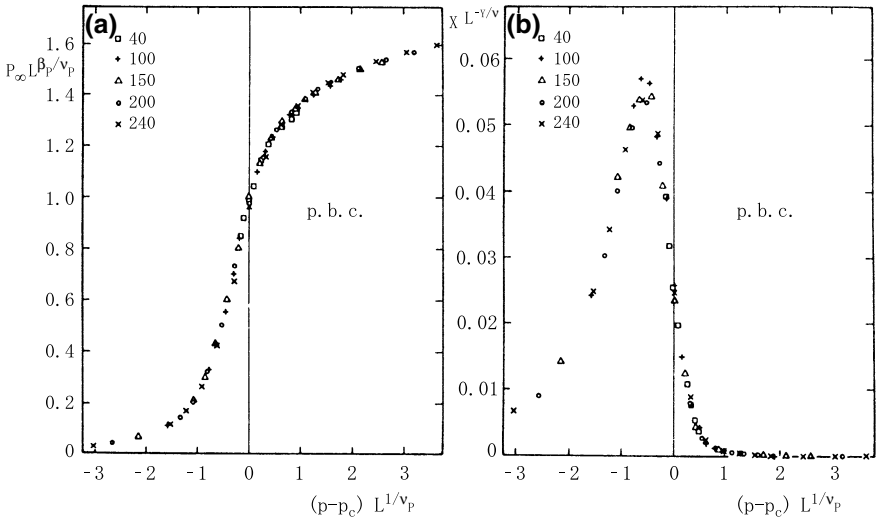


Fig. 2.10 Finite-size scaling plots of the data for the percolation probability $P_\infty^{(L)}(p)$ shown in Fig. 2.9a and for the percolation susceptibility (Fig. 2.9b, c). Here $L^{\beta_p/\nu_p} P_\infty^{(L)}(p)$ is plotted versus $(p - p_c)L^{1/\nu_p}$ (**a**), while $L^{-\gamma_p/\nu_p}\chi^{(L)}(p)$ is plotted versus $(p - p_c)L^{1/\nu_p}$ (**b**), with the following choice of exponents: $\beta_p = 0.139$, $\nu_p = \frac{4}{3}$, $\gamma_p = 2.41$. (From [2.63])

$$P_\infty^{(L)}(p) = L^{-\beta_p/\nu_p}\tilde{P}\left(L/\xi_p\right),$$
$$\tilde{P}(x \gg 1) \propto x^{\beta_p/\nu_p} \propto L^{\beta_p/\nu_p}\,(p - p_c)^{\beta_p}, \tag{2.63}$$

which expresses the principle that the effects of the finite size on the percolation transition are controlled by the ratio of the correlation length and the lattice linear dimension. Figure 2.10a shows a test of this relation [2.63]. A similar relation can be postulated for the susceptibility $\chi^{(L)}(p)$:

$$\chi^{(L)}(p) = L^{\gamma_p/\nu_p}\tilde{\chi}(L/\xi_p), \quad \tilde{\chi}(x) \underset{x\to\infty}{\propto} x^{-\gamma_p/\nu_p}. \tag{2.64}$$

The asymptotic behavior of the scaling functions $\tilde{P}(x)$, $\tilde{\chi}(x)$ for large x always follows from the condition that in the thermodynamic limit all powers of L must cancel out and the correct power laws as written in (2.61) and (2.62) must result. Figure 2.10b shows a test of the finite-size scaling relation (2.64) for the percolation susceptibility [2.63].

The fact that for p near p_c we must have $\chi^{(L)}(p \approx p_c) \propto L^{\gamma_p/\nu_p}$ can be obtained directly from the definition of χ and (2.62a), noting that the sum must be cut off at cluster sizes of the order of l_L:

$$\begin{aligned}\chi^{(L)} &\cong \frac{1}{p}\sum_{l=1}^{l_L} l^2 n_l(p) \underset{p \text{ near } p_c}{\cong} \tilde{n}(0)\sum_{l=1}^{l_L} l^{2-\tau}/p_c \approx \tilde{n}(0)p_c^{-1}\int_0^{l_L} \mathrm{d}l\, l^{2-\tau} \\ &= \tilde{n}(0)p_c^{-1} l_L^{3-\tau}/(3-\tau) \propto L^{d_f(3-\tau)} \\ &= L^{d(1-1/\delta_p)/(1+1/\delta_p)} = L^{\gamma_p/\nu_p}. \end{aligned} \tag{2.65}$$

The self-consistency of this geometric interpretation of finite-size scaling at the percolation transition is noted from calculating the characteristic length ξ_p from (2.60) and (2.61) as

$$\begin{aligned}\chi\xi_p^2 &= \sum_{l=1}^{\infty} r_l^2 l^2 n_l(p) \cong \hat{r}^2 \int_0^{\infty} l^{2/d_f+2-\tau}\tilde{n}\left[l^{\sigma}\left(1-\frac{p}{p_c}\right)\right]\mathrm{d}l \\ &= \left(1-\frac{p}{p_c}\right)^{-(3-\tau+2/d_f)/\sigma} \hat{r}^2 \int_0^{\infty} x^{2/d_f+2-\tau}\tilde{n}\,(x^{\sigma})\,\mathrm{d}x \\ &\propto \left(1-\frac{p}{p_c}\right)^{-2\nu_p-\gamma_p}, \end{aligned} \tag{2.66}$$

since $1/\sigma d_f = \nu_p$.

When one wishes to use (2.63) and (2.64) to locate p_c and determine the exponents β_p/ν_p, $1/\nu_p$ (or γ_p/ν_p), one can try a simultaneous best-fitting procedure such that the family of curves $P_\infty^{(L)}(p)$ collapses onto a single curve, the scaling function $\tilde{P}(x)$, as well as possible [or that the family of curves $\chi^{(L)}(p)$ collapses onto a single curve, the scaling function $\tilde{\chi}(x)$, when plotted appropriately]. Figure 2.10 shows examples of this sort of "data collapsing". However, to obtain accurate estimates it is important to have rather large values of L, since (2.63) and (2.64) hold only asymptotically in the limit $L \to \infty$: for finite L, there are systematic corrections to finite-size scaling which lead to systematic deviations from perfect data collapsing. Since these correction terms are completely disregarded in plots such as shown in Fig. 2.10, one can never be sure whether the "best-fit" values for p_c, β_p/ν_p, $1/\nu_p$, γ_p/ν_p are systematically offset from their true values. This problem does not occur when one samples the probability $P_s(p)$ that a "spanning cluster" (reaching from one boundary to the opposite one) occurs in the system. Since for the infinite lattice $P_s(p) = 0$ for $p < p_c$ and $P_s(p) = 1$ for $p \geq p_c$, the finite-size scaling relation for $P_s(p)$ simply reads

$$P_s^{(L)}(p) = \tilde{P}_s(L/\xi_p). \tag{2.67}$$

As a consequence, different curves $P_s^{(L)}(p)$ for different choices of L should intersect at $p = p_c$ in a common intersection point $\tilde{P}_s(0)$, apart from corrections to scaling. Extrapolation of these intersection points to the limit $L \to \infty$ hence yields an estimate for p_c which is not biased by the choice of the critical exponents and should be free from the systematic errors noted above. Figure 2.11 shows an example demonstrating that accurate estimates are easily obtained from this intersection

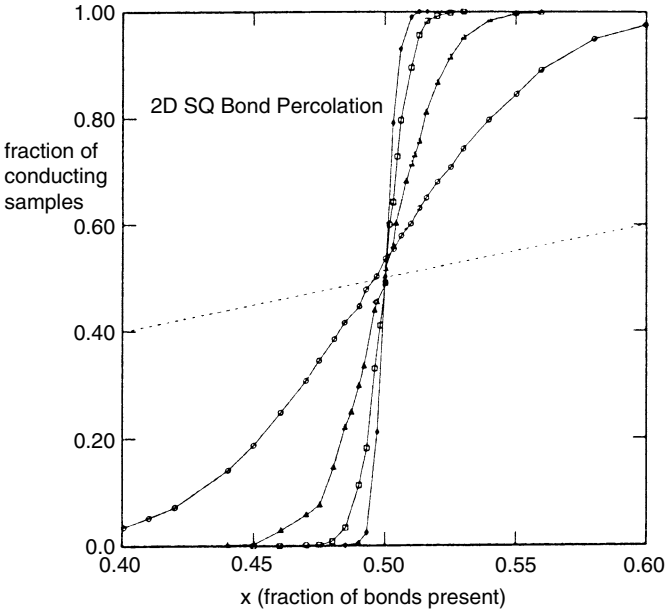


Fig. 2.11 Fraction of conducting samples in the bond percolation problem on the square lattice where a bond is conducting with probability $p(= x)$ and isolating with probability $1 - p$, plotted as a function of $p(= x)$ for various linear dimensions L: (*circles*) $L = 16$; (*triangles*) $L = 64$; (*squares*) $L = 200$; (*diamonds*) $L = 512$. A sample is conducting only if a "spanning cluster" occurs, and hence this quantity just measures $P_{\mathrm{s}}^{(L)}(p)$. The percolation threshold p_{c} is obtained as the common intersection point of these curves. (From [2.64])

method [2.64]. The reader who works through Exercises 3.18, 3.25 of Sect. 3.2.4 will generate similar Monte Carlo data.

2.3.3 Broken Symmetry and Finite-Size Effects at Thermal Phase Transitions

Effects of finite size on the percolation transition are relatively easy to understand, firstly, because of their obvious geometric interpretation, secondly, because the situation is simple due to the lack of any spontaneous symmetry breaking at the transition. We now discuss ordinary, thermally driven phase transitions, where the system state changes from a disordered state at high temperatures to a spontaneously ordered state at temperatures lower than some critical temperature T_{c} of a second-order phase transition. The prototype example is again the Ising ferromagnet (2.1), where this low-temperature phase is a state with nonzero spontaneous magnetization ($\pm|M_{\mathrm{sp}}|$ for zero applied field H). It is well known, of course, that this spontaneous

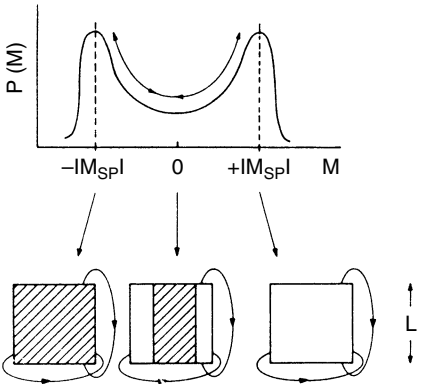


Fig. 2.12 Probability distribution $P(M)$ for a finite system at a temperature T less than the critical temperature T_c where in the thermodynamic limit a spontaneous magnetization $\pm|M_{sp}|$ appears. While $P(M)$ has a peak near $M = -|M_{sp}|$, where the finite lattice has a more or less uniformly negative magnetization (*shaded*), and a second peak where it has a more or less uniformly positive magnetization (*unshaded*), in between $P(M)$ has a minimum which is nonzero even for $M = 0$. This state corresponds to a nonuniform distribution of magnetization in the finite lattice: making use of the periodic boundary conditions, two domains of opposite magnetization coexist. In a finite lattice, the system spontaneously makes excursions from states with uniformly negative magnetization through this intermediate mixed-phase state to states with uniformly positive magnetization, and vice versa

symmetry breaking can occur in the thermodynamic limit only. For a finite system, there is always a nonzero probability that the system may pass from a state near $+|M_{sp}|$ to a state near $-|M_{sp}|$, as well as in the opposite direction, see Fig. 2.12. Therefore, the magnetization M at zero field vanishes for all nonzero temperatures,

$$M(T, H = 0) = \frac{1}{N} \sum_{i=1}^{N} \langle S_i \rangle_{T,H=0} = 0, \tag{2.68}$$

irrespective of the value of N. In fact, the first-principles definition of the order parameter M_{sp} considers $M(T, H)$, first taking the thermodynamic limit $N \to \infty$, and afterwards letting H tend to zero,

$$M_{sp} = \lim_{H \to 0} \lim_{N \to \infty} M(T, H). \tag{2.69}$$

This double limiting procedure is rather inconvenient to use in practical Monte Carlo work, and hence is applied only in exceptional cases, such as spin glasses, where the ordering is very hard to understand [2.65]. In Ising ferromagnets below T_c but not too close to the critical point, one finds that a magnetization $+M$ (or $-M$, depending on the initial condition) is sufficiently metastable for long observation times, and hence estimates with reasonable accuracy can be obtained although one does not sample the full equilibrium ensemble [2.66]. However, even above T_c a small magnetization

will typically be found, due to fluctuations which in a finite system observed over a finite time have not completely averaged out. One will find a value $\pm\delta M$, where δM depends on both the size of the system and the observation time t_{obs}. Similarly, below T_{c} the magnetization fluctuates in the range $M \pm \delta M$ or $-M \mp \delta M$, and one cannot make δM arbitrarily small by making t_{obs} larger and larger. If t_{obs} becomes of the order of the ergodic time t_{e}, which is the time needed to observe transitions from $+M$ to $-M$ or vice versa, one would start averaging the magnetization to zero. This situation becomes particularly cumbersome near T_{c}, where M itself strongly decreases, while δM increases until the fluctuations become comparable with the order parameter itself. The reader is urged to work through Sect. 3.4.1 and to program Exercise 3.36 to verify these remarks.

To avoid these problems, the standard recipe is to record the root mean square order parameter [2.67]

$$M_{\mathrm{rms}} = \sqrt{\langle M^2 \rangle_T} = \left\langle \left(\sum_{i=1}^{N} S_i / N \right)^2 \right\rangle_T^{1/2} = \frac{1}{N} \left(\sum_{i,j=1}^{N} \langle S_i S_j \rangle_T \right)^{1/2} . \tag{2.70}$$

In particular, (2.70) must be used for isotropic spin systems, (2.2) and (2.3), where one has a vector order parameter $\boldsymbol{M}_{\mathrm{sp}}$, whose orientation is not even metastable. One observes a sort of "rotational diffusion" of the unit vector along $\boldsymbol{M}_{\mathrm{sp}}$ [2.68] and hence each component of $\langle \boldsymbol{M}_{\mathrm{sp}} \rangle_T$ steadily decreases with increasing observation time.

Of course, in a finite system the order parameter M_{rms} is nonzero at all temperatures. Even at infinite temperatures, where $\langle S_i S_j \rangle = \delta_{ij}$, one still obtains $M_{\mathrm{rms}} = 1/\sqrt{N}$ [2.67]. At temperatures slightly above T_{c}, where the correlation function $G(\boldsymbol{r}_{ij}) \equiv \langle S_i S_j \rangle_T$ is slowly decaying, M_{rms} is even much larger. With periodic boundary conditions $G(\boldsymbol{r}_{ij})$ is translationally invariant and hence $M_{\mathrm{rms}} = (\sum_{i=1}^{N} \langle S_i S_j \rangle_T / N)^{1/2}$. Let us now discuss this expression at T_{c} itself [2.21]. In an infinite system, the decay of correlations at T_{c} is described by

$$G(\boldsymbol{r}_{ij}) \underset{|\boldsymbol{r}_{ij}| \to \infty}{\longrightarrow} \hat{G} |\boldsymbol{r}_{ij}|^{-(d-2+\eta)}, \tag{2.71}$$

where $\hat{G}$ is a "critical amplitude" and η a critical exponent. We now approximate $G(\boldsymbol{r}_{ij})$ in the finite system of size L^d by taking $G(\boldsymbol{r}_{ij})$ also from (2.71) if $|\boldsymbol{r}_{ij}| < L/2$. Then we obtain [2.21]

$$\sum_{i=1}^{N}\langle S_i S_j\rangle_T \propto \int_0^{L/2} r_{ij}^{d-1}\mathrm{d}r_{ij}\,\langle S_i S_j\rangle_T$$
$$\propto \int_0^{L/2} r_{ij}^{1-\eta}\mathrm{d}r_{ij} \propto L^{2-\eta}, \tag{2.72a}$$

and hence M_{rms} becomes ($N = L^d$)

$$M_{\mathrm{rms}}^{T=T_c} \propto \left(L^{2-d-\eta}\right)^{1/2} \propto L^{-\beta/\nu}, \tag{2.72b}$$

where we have used the scaling laws $(2-\eta) = \gamma/\nu$, $\mathrm{d}\nu = 2\beta + \gamma$ [2.69]. Note the similarity of (2.72b) to the corresponding result for the percolation problem, $P_\infty^{(L)}(p_c) \propto L^{-\beta_p/\nu_p}$.

2.3.4 *The Order Parameter Probability Distribution and Its Use to Justify Finite-Size Scaling and Phenomenological Renormalization*

Equation (2.72b) is already a result characteristic of finite-size scaling theory [2.23–2.26]. To describe this theory more systematically, we now discuss the probability distribution of the order parameter s [2.25]. For $T > T_c$ and linear dimensions L exceeding the correlation length ξ of order parameter fluctuations ($\xi \propto |T - T_c|^{-\nu}$), this distribution should be Gaussian as considered in (2.33)

$$P_L(s) = L^{d/2}\left(2\pi k_B T \chi^{(L)}\right)^{1/2} \exp\left[-s^2 L^d / \left(2k_B T \chi^{(L)}\right)\right], \tag{2.73}$$
$$T > T_c, \quad H = 0.$$

The "susceptibility" $\chi^{(L)}$ defined in (2.73) from the half-width of the distribution should smoothly tend towards the susceptibility χ of the infinite system as $L \to \infty$ (remember $\chi \propto |T - T_c|^{-\gamma}$). For $T < T_c$ but again $L \gg \xi$, the distribution is peaked at values $\pm M_L$ near $\pm M_{sp}$; near these peaks again a description in terms of Gaussians applies (while a different behavior occurs near $s = 0$ [2.25, 2.70])

$$P_L(s) = \frac{L^{d/2}}{\left(2\pi k_B T \chi^{(L)}\right)^{1/2}}\left[\frac{1}{2}\exp\left(-\frac{(s - M_L)^2 L^d}{2k_B T \chi^{(L)}}\right)\right.$$
$$\left. + \frac{1}{2}\exp\left(-\frac{(s + M_L)^2 L^d}{2k_B T \chi^{(L)}}\right)\right], \quad T < T_c, \quad H = 0. \tag{2.74}$$

The small value of $P_L(s \approx 0) \propto \exp(-2L^{d-1} f_{\mathrm{int}}/k_B T)$, where f_{int} is the interfacial tension of the system [2.70], measures the probability that the system moves away from the region near $+M_L$ by the spontaneous creation of two interfaces of

size L^{d-1} to form a domain of negative magnetization spanning the system. By this mechanism the system can pass from $+M_L$ to $-M_L$ and vice versa, see Fig. 2.12. Since the observation time needed to observe these transitions increases with L as $P_L(s = M_L)/P_L(s = 0)$, for large L such transitions cannot be seen during reasonable observation times. Thus one does not sample the full symmetric distribution (2.74) for which $\langle s\rangle_L = \int_{-\infty}^{+\infty} s\,P_L(s)\mathrm{d}s = 0$, but rather only one half of it:

$$\langle s\rangle'_L = \frac{\displaystyle\int_0^\infty s\,P_L(s)\,\mathrm{d}s}{\displaystyle\int_0^\infty P_L(s)\,\mathrm{d}s} = \langle|s|\rangle_L. \tag{2.75}$$

When (2.74) is an accurate description of the actual distribution $P_L(s)$, the restricted average $\langle s\rangle'_L$ coincides with the peak position M_L, but due to deviations from the Gaussian distribution in an actual simulation these quantities may differ from each other. However, when extrapolated to the thermodynamic limit, all these quantities should yield the spontaneous magnetization

$$\lim_{L\to\infty} M_L = \lim_{L\to\infty} \langle|s|\rangle_L = \lim_{L\to\infty} \langle s^2\rangle_L^{1/2} = M_{\mathrm{sp}}. \tag{2.76}$$

Of course, these relations are more convenient than using (2.69). Figure 2.13 illustrates the use of these relations, where only a single system of size $N = 24^3$ was simulated but the order parameter distribution in *subsystems* of different linear dimensions $L < 24$ was recorded, for several values of L simultaneously in one run [2.25, 2.71]. Similarly, the susceptibility can be estimated both from the fluctuation-dissipation relation (relating it to magnetization fluctuations), and the half-widths Δs, or the heights $P_L(0)$, $P_L(M_L)$ of the peaks:

$$\begin{aligned}\lim_{L\to\infty} \frac{\langle s^2\rangle L^d}{k_{\mathrm{B}}T} &= \lim_{L\to\infty} \frac{P_L^{-2}(0)L^d}{2\pi k_{\mathrm{B}}T} \\ &= \lim_{L\to\infty} \frac{(\Delta s)^2 L^d}{8k_{\mathrm{B}}T\ln 2} = \chi, \quad T > T_{\mathrm{c}},\end{aligned} \tag{2.77a}$$

$$\begin{aligned}\lim_{L\to\infty} \frac{\langle s^2\rangle - \langle|s|\rangle^2}{k_{\mathrm{B}}T} L^d &= \lim_{L\to\infty} \frac{P_L^{-2}(M_L)L^d}{8\pi k_{\mathrm{B}}T} \\ &= \lim_{L\to\infty} \frac{(\Delta s)^2 L^d}{8k_{\mathrm{B}}T\ln 2} = \chi, \quad T < T_{\mathrm{c}}.\end{aligned} \tag{2.77b}$$

Equations (2.73) and (2.74) hold for $L \gg \xi$. In a practical calculation ξ is usually not known, then the Gaussian character of the distribution is conveniently studied by calculating the fourth-order cumulant U_L [2.25] (or the equivalent quantity $g_L = -3U_L$ called the renormalized coupling constant [2.72])

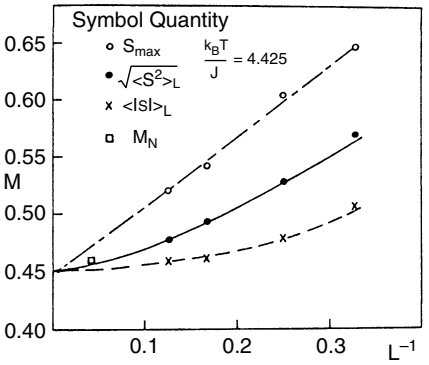


Fig. 2.13 Estimates of the spontaneous magnetization of the three-dimensional Ising model with nearest-neighbor interaction on the simple cubic lattice at a temperature ($k_BT/J = 4.425$) below criticality ($k_BT_c/J \approx 4.51$), where J is the exchange constant, see (2.1), as obtained from extrapolating the size dependence of the position of the maximum (s_{max}) and the moments $\langle s^2\rangle_L$ and $\langle|s|\rangle_L$ towards $L^{-1} \to 0$, for subsystems of a system of size 24^3. The direct estimate for the magnetization of the total system (M_N) is also included. (From [2.71])

$$U_L = 1 - \frac{\langle s^4\rangle_L}{3\langle s^2\rangle_L^2}. \tag{2.78}$$

For $T > T_c$ and $L \gg \xi$, one can show that U_L decreases towards zero as $U_L \propto L^{-d}$ [2.25]. For $T < T_c$ and $L \gg \xi$, one can show from (2.74) that U_L tends to $U_\infty = 2/3$. For $L \ll \xi$, on the other hand, U_L varies only weakly with temperature and linear dimension, it stays close to the (universal but nontrivial) "fixed-point" value U^*.

This behavior of the cumulant makes it very useful for obtaining estimates of T_c itself which are not biased by any assumptions about critical exponents [2.25]. One may plot U_L versus T for various L's and estimate T_c from the common intersection point of these curves. As shown below, finite-size scaling implies the existence of such a common intersection point. Due to corrections to finite-size scaling, there may be some scatter in the intersection points for different pairs (L, L') if one works with very small linear dimension (Fig. 2.14). Nevertheless, the accuracy of this "cumulant intersection method" is quite good. It has been applied to a variety of models [2.13, 2.25, 2.29, 2.71–2.78] including sophisticated examples such as spin glasses [2.77] and XY models with cubic anisotropy [2.73]. Note also that the estimate for the critical point of the $3d$ Ising model obtained by this technique on the Santa Barbara special purpose computer [2.72], $J/k_BT_c = 0.221650(\pm 5)$, is competitive in accuracy with the most extensive series expansion and Monte Carlo renormalization group estimates [2.79] available, although the work in [2.72] is hampered by some spurious finite-size effects due to pseudorandom numbers of bad quality, as shown in related work done on different computers [2.80–2.82].

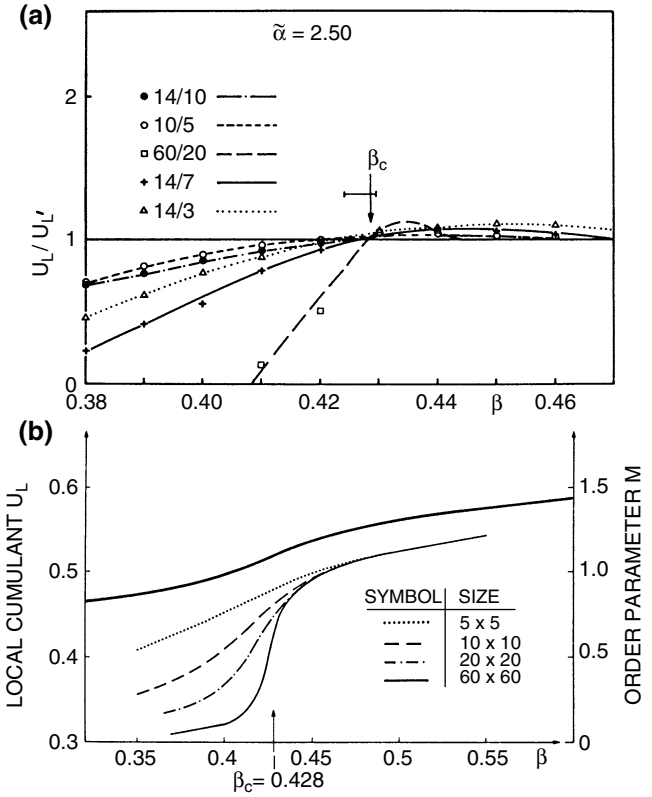


Fig. 2.14 **a** Determination of the critical coupling $\tilde{\beta}_c$ of the ϕ^4 model on the square lattice (2.43), where parameters are renormalized as $\tilde{a} = (A + 2dC)^2/B$, $\tilde{\beta} = -C(A + 2dC)/B$, for the particular choice $\tilde{a} = 2.50$, and for various ratios $U_L/U_{L'}$, for pairs (L, L') as indicated in the figure. Note the large scale of the abscissa to demonstrate the accuracy of this method. **b** Variation of the order parameter M_L defined as $M_L = \langle |\sum_{i=1}^{N} \phi_i / \sqrt{-(A + 2dC)/B}|/N \rangle$ for the same model. (From [2.13])

We now turn to the description of the distribution function $P_L(s)$ in the region where L is not necessarily larger than ξ. The key idea is that $P_L(s)$, which is a function of L, s, and ξ (which expresses the temperature dependence via $\xi \propto |1 - T/T_c|^{-\nu}$, as noted above), does not depend separately on these three variables, but only on two scaled combinations, L/ξ and $s\xi^{\beta/\nu}$ [2.25]:

$$P_L(s) = \xi^{\beta/\nu} P\left(L/\xi, s\xi^{\beta/\nu}\right) = L^{\beta/\nu} \tilde{P}\left(L/\xi, sL^{\beta/\nu}\right). \tag{2.79}$$

Here the power-law prefactors are needed to ensure the normalization,

$$\int_{-\infty}^{+\infty} \mathrm{d}s\, P_L(s) = 1,$$

and in the last part of (2.79) we have used

$$sL^{\beta/\nu} = \left(s\xi^{\beta/\nu}\right)(L/\xi)^{\beta/\nu}$$

instead of $s\xi^{\beta/\nu}$ as our second scaling variable. From (2.79), which should hold in the limit $L \to \infty, \xi \to \infty$, but L/ξ finite, it is straightforward to derive the standard finite-size relations [2.23, 2.24] analogous to (2.63), (2.64), and (2.67) by taking suitable moments of the distribution,

$$\langle |s| \rangle_L = L^{-\beta/\nu} \tilde{M}(L/\xi), \tag{2.80a}$$

$$\chi'(L,T) \equiv L^d \left(\langle s^2 \rangle_L - \langle |s| \rangle_L^2 \right) / k_\mathrm{B} T = L^{\gamma/\nu} \tilde{\chi}(L/\xi), \tag{2.80b}$$

$$U_L = 1 - \frac{\tilde{\chi}_4(L/\xi)}{3\left[\tilde{\chi}_2(L/\xi)\right]^2}. \tag{2.80c}$$

Here $\tilde{\chi}_4(L/\xi)$ is related to the 4th moment, $\langle s^4 \rangle_L \equiv L^{-4\beta/\nu} \tilde{\chi}_4(L/\xi)$, and $\tilde{\chi}_2(L/\xi)$ to the second, $\langle s^2 \rangle_L \equiv L^{-2\beta/\nu} \tilde{\chi}_2(L/\xi)$. In (2.80b) we have defined a function $\chi'(L,T)$ which for $T < T_\mathrm{c}$ tends to the standard susceptibility χ in the thermodynamic limit (where $\chi \equiv \lim_{H\to 0} \lim_{L\to\infty} \partial \langle s \rangle_L / \partial H$), as is obvious from (2.77b), *while for $T > T_\mathrm{c}$ it does not*. In fact, from (2.73) it is straightforward to show that for $T > T_\mathrm{c}$ (remember $\lim_{L\to\infty} \chi^{(L)} = \chi$)

$$\lim_{L\to\infty} \chi'(L,T) = \lim_{L\to\infty} L^d \left(\langle s^2 \rangle_L - \langle |s| \rangle_L^2 \right) / k_\mathrm{B} T = \chi \left(1 - \frac{2}{\pi} \right). \tag{2.80d}$$

Thus χ' diverges with the same exponent as χ, but the critical amplitude is reduced by a factor $1 - 2/\pi$. This point is often confused in the literature, where sometimes $\lim_{L\to\infty} \chi'(L,T)$ is taken as an estimate for χ above T_c, and sometimes the formula $\chi = \lim_{L\to\infty} L^d \left(\langle s^2 \rangle_L - \langle s \rangle_L^2 \right) / k_\mathrm{B} T$ is used at all temperatures. This latter formula, however, in the absence of a symmetry-breaking field makes sense only for $T > T_\mathrm{c}$, where $\langle s \rangle_L^2 \equiv 0$ can be used, and then this formula reduces to (2.77a). For $T < T_\mathrm{c}$, on the other hand, in a Monte Carlo simulation for a finite system near T_c, $\langle s \rangle_L^2$ is not a useful quantity. For observation times which are much smaller than the ergodic time τ_e, $\langle s \rangle_L^2 \approx \langle |s| \rangle_L^2$, which is close to the squared spontaneous magnetization M_sp^2 (2.76), while for observation times which are much larger than τ_e, $\langle s \rangle_L^2 \equiv 0$, and for observation times of the order of τ_e one may get a rather erratic behavior where $\langle s \rangle_L^2$ can take any value in between zero and $\langle |s| \rangle_L^2$, when the magnetization has jumped back and forth between $+\langle |s| \rangle_L$ and $-\langle |s| \rangle_L$ just a few times. Now τ_e is of order τ_max at T_c, see (2.85)–(2.88) below, and much larger than τ_max for T much less than T_c, with τ_max being the maximum "intrinsic" relaxation time. Note that for the magnetization, τ_max and the time τ_A considered in (2.54)–(2.72) are of the same order of magnitude. Since we have to work with an observation time $t_\mathrm{obs} \gg \tau_A \approx \tau_\mathrm{max}$, below T_c we will necessarily encounter a temperature where $t_\mathrm{obs} \equiv \tau_\mathrm{e}$, and hence $\langle s \rangle_L^2$ is meaningless there. Thus it is an inevitable consequence of symmetry breaking that above T_c and below T_c in finite systems *different fluctuation* formulae must be

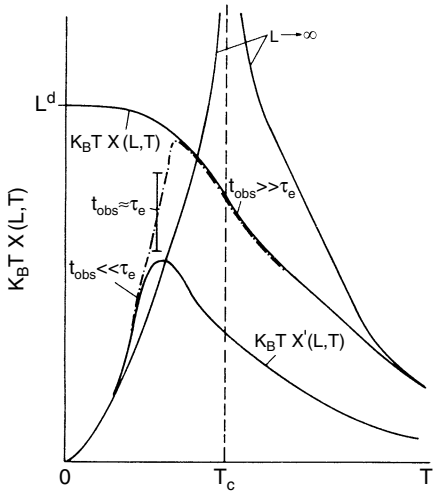


Fig. 2.15 Schematic temperature variation of the normalized susceptibilities $k_B T \chi(L, T)$ and $k_B T \chi'(L, T)$ as defined in (2.80b) and via $k_B T \chi(L, T) \equiv L^d \langle s^2 \rangle_L$. For $T \to 0$ there are no fluctuations, thus $\langle s^2 \rangle = \langle |s| \rangle^2 = 1$ and hence $k_B T \chi'(L, T \to 0) \to 0$ while $k_B T \chi(L, T \to 0) \to L^d$. Outside the regime of the finite-size rounding, $\chi(L, T)$ tends to the standard susceptibility χ for $T \geq T_c$, while for $T < T_c$ it is $\chi'(L, T)$ that tends to χ. While the divergence of χ is rounded off to a finite peak in the function $T\chi'(L, T)$, the function $T\chi(L, T)$ is monotonically increasing for decreasing T due to the onset of the order parameter. The *dash-dotted curve* illustrates the standard susceptibility formula $L^d(\langle s^2 \rangle_L - \langle s \rangle_L^2)$, which, unlike (2.80b), does not involve any absolute value: for zero field this expression is not well defined for the temperature region for which the observation time t_{obs} is of the same order as the ergodic time τ_e, where erratic fluctuations therefore occur. For $L \to \infty$, this temperature region shrinks and ultimately coincides with T_c

used to extract the susceptibility of the infinite system with the correct prefactor, as written down in (2.77a) and (2.77b). This behavior of the susceptibilities $\chi(L, T)$ and $\chi'(L, T)$ is illustrated schematically in Fig. 2.15.

One immediately finds that the order parameter at criticality behaves as $\langle |s| \rangle_L \propto L^{-\beta/\nu}$, consistent with our previous result (2.72b). The "fixed point" value of the cumulant is now interpreted as $U^* = 1 - \tilde{\chi}_4(0)/3[\tilde{\chi}_2(0)]^2$.

As was the case for (2.63) and (2.64), for the percolation problem, (2.80a)–(2.80c) form the basis for a study of critical properties of the infinite system, extracted from the Monte Carlo results for finite lattices. Figure 2.16 shows another example of data collapsing [2.13, 2.25, 2.27, 2.37, 2.71, 2.76, 2.82, 2.83]. Studying, for instance, the order parameter $M_L = \langle |s| \rangle_L$ for various L's, we have a family of curves (Fig. 2.14b). Multiplying M_L by a factor $L^{\beta/\nu}$ and the reduced temperature $|1 - T/T_c| = |1 - \tilde{\beta}_c/\tilde{\beta}|$ by a factor $L^{1/\nu}$, the family of curves should collapse onto the two branches (for $T > T_c$ and $T < T_c$, respectively) of the scaling function $\tilde{M}$. The figure already demonstrates the disadvantage of this method: one simultaneously has to fit three parameters $\{T_c,\ 1/\nu,\ \beta/\nu\}$, but since one often includes Monte Carlo "data" for which neither L nor ξ are very large, there are systematic corrections

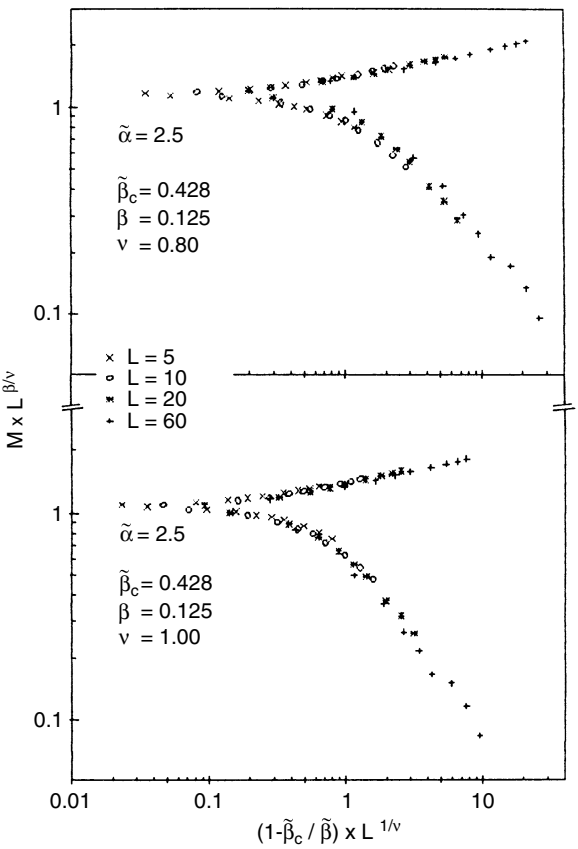


Fig. 2.16 Plot of $M_L L^{\beta/\nu}$ versus $(1 - \tilde{\beta}_c/\tilde{\beta})L^{1/\nu}$ for the ϕ^4 model on the square lattice with the parameter choice $\tilde{\alpha} = 2.5$, $\beta_c = 0.428$ (for the definition of parameters see Fig. 2.14) and $L = 5$, 10, 20, and 60. Finite blocks with periodic boundary conditions were used throughout, as in Fig. 2.14. *Upper part*, the choice of exponents $\beta = 0.125$, $\nu = 0.80$. *Lower part*, the standard Ising choice ($\beta = 0.125$, $\nu = 1.00$). (From [2.13])

to the asymptotic finite-size scaling expressions, and complete superposition of the curves is prevented. Moreover, there is not really a unique "fit", and hence the actual accuracy of the method is somewhat hard to ascertain. Nevertheless it has often yielded useful results (e.g., [2.83, 2.84]).

Alternative phenomenological renormalization methods based on (2.80) derive critical exponents from relations such as ($L' = bL$ with a scale factor $b > 1$) [2.25]

$$\frac{1}{\nu} = \left.\frac{\ln\left(\partial U_{bL}/\partial U_L\right)}{\ln b}\right|_{U^*},$$
$$\frac{\gamma}{\nu} = \frac{\ln\left[\chi(bL, T_c)/\chi(L, T_c)\right]}{\ln b}. \tag{2.81}$$

One advantage of this method is that T_c and the exponents $1/\nu$, γ/ν are estimated independently of each other; another advantage is that one can analyze systematic errors due to corrections to finite-size scaling. At T_c such corrections are, to leading order, additive terms with less strongly divergent exponents

$$\chi(L, T_c) = L^{\gamma/\nu}\tilde{\chi}(0)(1 + \chi^{\text{corr}} L^{-x_{\text{corr}}} + \cdots), \tag{2.82}$$

χ^{corr} being another amplitude factor and x_{corr} the leading correction exponent. Now (2.81) is replaced by

$$\frac{\ln[\chi(bL, T_c)/\chi(L, T_c)]}{\ln b} = \frac{\gamma}{\nu} - \frac{\chi^{\text{corr}} L^{-x_{\text{corr}}}}{\ln b}(1 - b^{-x_{\text{corr}}}) + \cdots. \tag{2.83}$$

Thus, plotting estimates for γ/ν (or $2\beta/\nu$, cf. [2.25, 2.71]) versus $1/\ln b$, one obtains for each L a different curve, which for $(\ln b)^{-1} \to 0$ must extrapolate linearly to the same value of γ/ν. The disadvantage, however, of this method is that extremely good statistical accuracy of the results for $\chi(L, T_c)$ is required, otherwise the procedure is not applicable. Landau and Binder [2.75] have applied this technique to a nontrivial two-dimensional example, the Ising antiferromagnet with nearest and next-nearest neighbor interactions J_{nn}, J_{nnn} (Fig. 2.17), which exhibits a transition to a layered antiferromagnetic structure for $J_{nnn}/J_{nn} > \frac{1}{2}$, belonging to the universality class of the XY model with cubic anisotropy [2.85], with nonuniversal critical exponents. The exponent estimates which can be obtained are comparable in accuracy to results obtained from Monte Carlo renormalization group (MCRG) [2.86], Monte Carlo data collapsing [2.83], high-temperature series extrapolations [2.87] and transfer matrix calculations for finite strips [2.88], see Fig. 2.18. The method is definitely superior

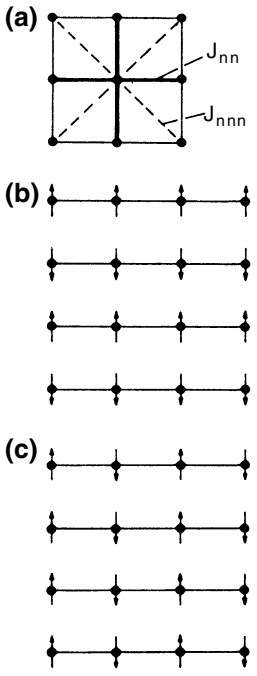


Fig. 2.17 **a** Square lattice with exchange J_{nn} between nearest neighbors (*full lines*) and exchange J_{nnn} between next nearest neighbors (*broken lines*). **b**, **c** Spin arrangements in the layered antiferromagnetic structure, where ferromagnetically aligned rows of up spins alternate with ferromagnetically aligned rows of down spins. These rows can be oriented in the x direction (**b**) or in the y direction (**c**). These two orderings are always degenerate with each other and hence should be considered as the two components of a two-component order parameter (ψ_x, ψ_y)

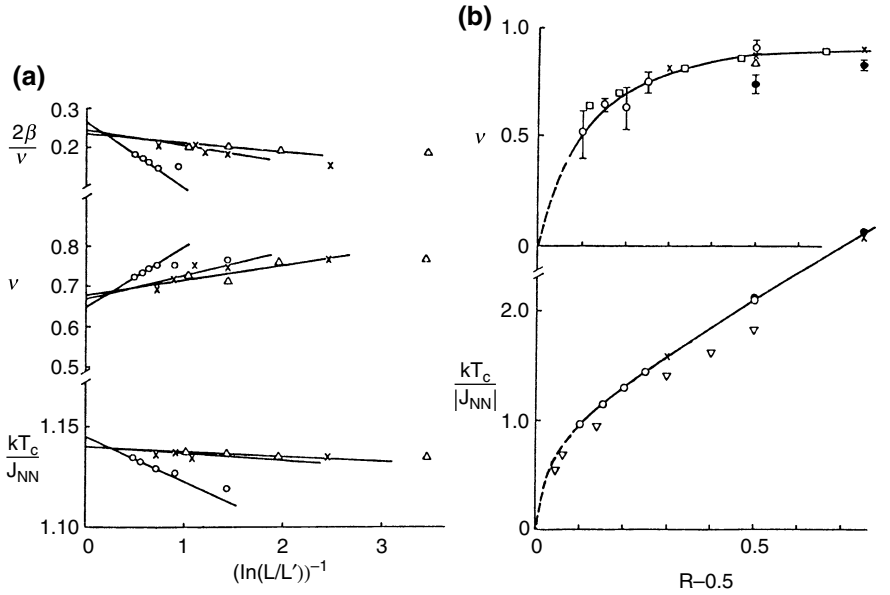


Fig. 2.18 **a** Variation of critical parameters with $\ln b = \ln(L/L')$ for the Ising antiferromagnet on the square lattice with $R \equiv J_{nnn}/J_{nn} = 0.65$. Data are for $L' = 4$ (○), $L' = 8$ (×), and $L' = 12$ (△); finite blocks with periodic boundary conditions were used throughout. (From [2.75]). **b** Variation of the correlation length exponent (*upper part*) and critical temperature T_c (*lower part*) with R. (*circles*) Results of phenomenological renormalization, as described in (**a**); (*triangles*) Monte Carlo "data collapsing" results of [2.83]; (*crosses*) MCRG results of [2.86]; (*filled circles*) series expansion results of [2.87]; (*squares*) transfer-matrix renormalization [2.88]; (*inverted triangles*) real-space renormalization group results [2.89]. (From [2.75])

to standard real-space renormalization group methods [2.89], but clearly involves a major effort in computing time. This is also true for the MCRG method, although it seems that there good results are obtained with somewhat smaller statistical effort, even in three dimensions [2.79]. This phenomenological renormalization method was also tried for the nearest-neighbor three-dimensional Ising model, using Monte Carlo results obtained with the Ising special purpose computer [2.72]. It was claimed [2.72] that the asymptotic critical regime where finite-size scaling holds is only reached for $L > 24$, but it is now clear [2.80–2.82] that this claim is wrong, due to inappropriate pseudorandom numbers used in [2.72], and the approach to the thermodynamic limit is very smooth also in the three-dimensional Ising model. Still, a highly accurate phenomenological renormalization study of the three-dimensional Ising model has not been done before 1991 [2.89a].

A variant of this phenomenological renormalization technique [2.90] avoids the use of the fourth-order cumulant U_L, (2.78) and (2.80d), and works with $\chi(L, T)$ only. One then forms the ratio considered in (2.83), namely the function $\varphi_{L,L'}(T)$

$$\varphi_{L,L'}(T) \equiv \frac{\ln\left[\chi(L',T)/\chi(L,T)\right]}{\ln(L'/L)} \tag{2.84}$$

for two pairs of sizes, (L, L') as well as (L', L''). The functions $\varphi_{L,L'}(T)$ and $\varphi_{L',L''}(T)$ should thus intersect at T_c, and the intersection point should yield γ/ν if corrections to finite-size scaling can be neglected [2.90]. This method has been applied successfully to the two-dimensional ANNNI model, where a transition to an incommensurate floating phase occurs [2.90].

At this stage, it is extremely helpful if the reader turns to Sect. 3.4.1 and works out Exercises 3.37, 3.38, and 3.41.

2.3.5 *Finite-Size Behavior of Relaxation Times*

So far, we have considered only static quantities and analyzed the respective finite-size effects, but also the critical singularity of the "intrinsic" relaxation time τ

$$\tau \propto \xi^z \propto |1 - T/T_c|^{-\nu z} \tag{2.85}$$

exhibits a finite size rounding, again controlled by the rule that the rounding sets in when L is comparable to ξ, and hence the maximum "intrinsic" relaxation time is

$$\tau_{\max} \propto L^z \quad (T = T_c). \tag{2.86}$$

From (2.56) and (2.86) we now realize why it is so difficult to obtain accurate results at T_c, where the statistical error of the magnetization takes on its maximum value

$$\begin{aligned} \langle(\delta M)^2\rangle_{T_c} &= \frac{2\tau_{\max}}{t_{\text{obs}}}\left(\langle M^2\rangle_{T_c} - \langle|M|\rangle_{T_c}^2\right) \\ &= \frac{2\tau_{\max}\chi'_{\max}k_B T_c}{t_{\text{obs}}L^d} \propto \frac{L^{z+\gamma/\nu-d}}{t_{\text{obs}}}. \end{aligned} \tag{2.87}$$

Since *on a serial computer* the CPU time for one Monte Carlo step per site increases as L^d, the CPU time needed to obtain data for $\langle|M|\rangle$ at T_c with a given constant accuracy increases as $L^{z+\gamma/\nu} \approx L^4$, independent of d for $d \leq 4$. Thus an increase by a factor of 10 in the linear dimension would require 10^4 more computing time! Thus we see why vector processors and parallel computers, for which there is less increase in CPU time needed for one Monte Carlo step per site with increasing linear dimension, are indispensable for the Monte Carlo simulation of very large lattices at T_c.

For $T \geq T_c$, the "intrinsic" order parameter relaxation time τ considered in (2.85) and (2.86) is in fact the absolutely largest relaxation time in the system. For $T < T_c$, this is not true due to the existence of the "ergodic" time τ_e. As was

mentioned after (2.74), this ergodic time is proportional to $P_L(s = M_L)/P_L(s = 0) \propto \exp(2L^{d-1} f_{\text{int}}/k_B T)$. Since near T_c, $f_{\text{int}}/k_B T \propto \xi^{-(d-1)}$, we can write a scaling law for the ergodic time as follows:

$$\tau_e \propto L^z \exp\left[const\,(L/\xi)^{d-1}\right], \quad T < T_c. \tag{2.88}$$

This relation has been verified by Miyashita and Takano [2.91], while (2.86) was studied in careful work by Wansleben and Landau [2.92].

Below T_c, in addition to the intrinsic relaxation time τ (2.85), which describes how a small deviation δM from the spontaneous magnetization M_{sp} relaxes, and to the ergodic time τ_e, describing the time needed to flip over the sign of the total magnetization, various nonlinear relaxation times of the magnetization are also of interest. If we start the system in a fully aligned ferromagnetic state, the relaxation times $\tau_M^{(\text{nl})}$, $\tau_E^{(\text{nl})}$ of the magnetization, energy, etc., needed to relax to their equilibrium values can be shown to diverge near T_c according to a power law different from (2.85) [2.93, 2.94]:

$$\tau_M^{(\text{nl})} \propto (1 - T/T_c)^{-(\nu z - \beta)}, \tag{2.89a}$$

$$\tau_E^{(\text{nl})} \propto (1 - T/T_c)^{-(\nu z - 1 + \alpha)}, \tag{2.89b}$$

where α is the critical exponent of the specific heat [2.69]. On the other hand, if we start the system in a disordered configuration and no symmetry breaking field is present, ordered domains form (Fig. 2.19) whose linear dimensions $l(t)$ grow according to a power law [2.5, 2.46]

$$l(t) = (\Omega t)^x, \quad \Omega \propto (1 - T/T_c)^{\nu z - \nu/x}. \tag{2.90}$$

Now with such a "disordered start" in the regime of the ordered phase the times to reach equilibrium are at least of the order of t_L defined by

$$l(t_L) \equiv L, \quad \text{i.e.,} \quad t_L = \Omega^{-1} L^{1/x} \propto \tau (L/\xi)^{1/x} \propto \xi^z (L/\xi)^{1/x}. \tag{2.91}$$

While the times $\tau_M^{(\text{nl})}$, $\tau_E^{(\text{nl})}$ are very small if one works at T far below T_c, the time t_L is still rather large since $x = \frac{1}{2}$ in the nonconserved kinetic Ising model [2.5, 2.46], and hence $t_L \propto L^2$. In addition, this time t_L is only a lower bound to establish true equilibrium: it only says that domains with sizes comparable to the system linear dimension are present. There is no guarantee that the system exists in a monodomain state. Often one finds [2.68, 2.95], however, that the system after the time t_L is in a state with a few frozen-in domain walls oriented parallel to a surface of the system, such that each wall is connected to itself via the periodic boundary condition. It then may take an impractically long time for the frozen-in domain walls to disappear from the system, particularly in three-dimensional systems at temperatures below the interface roughening transition temperature [2.96]. This consideration shows that for the Monte Carlo study of ordered phases it is advisable to choose an appropriate

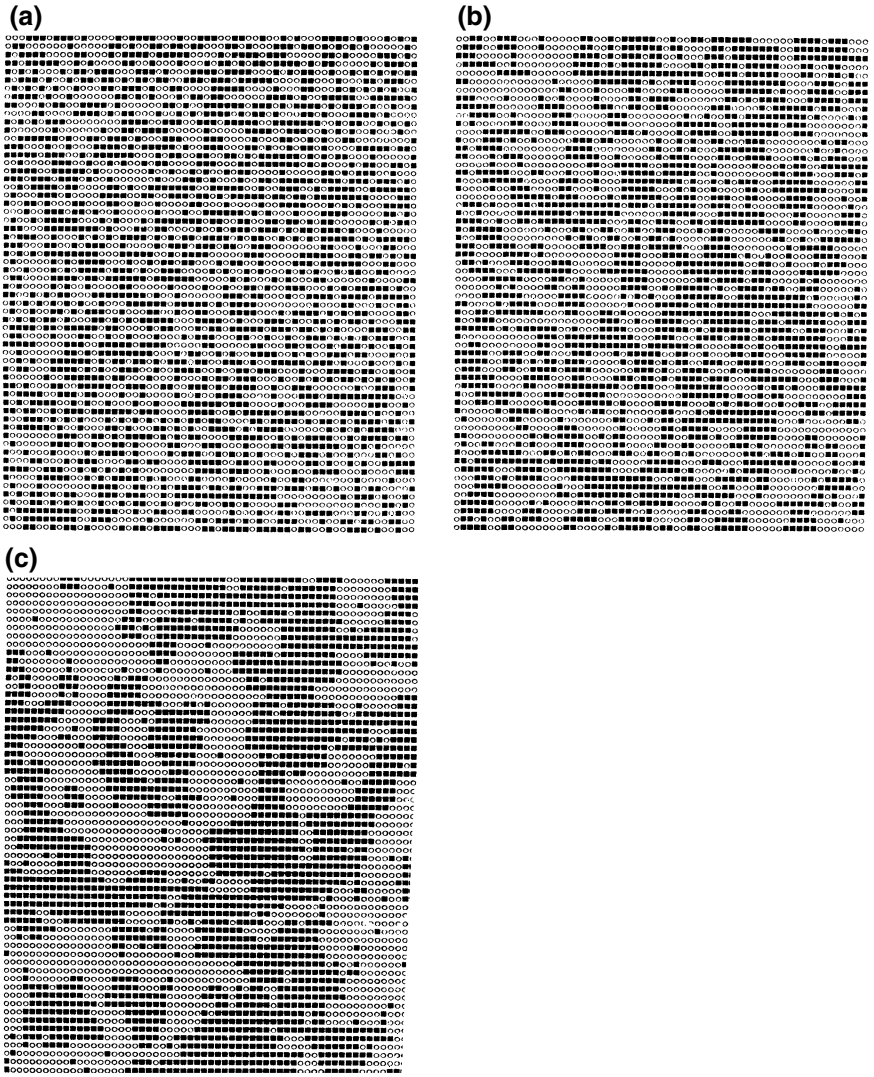


Fig. 2.19 Snapshot pictures of domains at various times after a start in a completely random initial configuration

initial condition, such as a fully ordered configuration. In systems where the detailed nature of the ordering is unknown, such as in spin glasses [2.41], a study of the low temperature phase is very difficult.

2.3.6 Finite-Size Scaling Without "Hyperscaling"

The finite-size scaling theory described so far, where one scales the linear dimension L with the correlation length ξ, rests on the validity of the hyperscaling relation between critical exponents, $d\nu = \gamma + 2\beta$ [2.25, 2.76]. Consequently, finite-size scaling in its standard form does not hold when hyperscaling is violated, which happens, for instance, in systems above their marginal dimensionality d^* where mean field predictions for critical exponents become valid (e.g., $d^* = 4$ for Ising models [2.69], $d^* = 6$ for percolation [2.60]). Then with fully periodic boundary conditions in an L^d geometry a simple modified form of finite-size scaling holds [2.76], the correlation length ξ being replaced by a *thermal* length l_T defined by [2.76]

$$l_\mathrm{T}^d = k_\mathrm{B} T \chi M^{-2} \propto |1 - T/T_\mathrm{c}|^{-(\gamma+2\beta)} = |1 - T/T_\mathrm{c}|^{-2}, \tag{2.92}$$

where in the last equality the mean field exponents $\gamma = 2\beta = 1$ were inserted. Equation (2.92) is understood by noting that (2.74) is true also for $d > d^*$, and the argument of the exponential function can be written as

$$\frac{(s \pm M)^2 L^d}{2k_\mathrm{B} T \chi} = \frac{(s/M \pm 1)^2}{2} \frac{L^d}{k_\mathrm{B} T \chi M^{-2}} = \frac{(s/M \pm 1)^2}{2} \left(\frac{L}{l_\mathrm{T}}\right)^d , \tag{2.93}$$

which shows that L scales with l_T. For $d < d^*$, $l_\mathrm{T} \propto \xi$ but for $d > d^*$, ξ has a different temperature dependence. The general scaling behavior replacing (2.79) is

$$\begin{aligned} P_L(s) &= (1 - T/T_\mathrm{c})^{-\beta} P' \left(L/l_\mathrm{T}, s(1 - T/T_\mathrm{c})^{-\beta}\right) \\ &= L^{d\beta/(\gamma+2\beta)} \tilde{P}' \left(L/l_\mathrm{T}, sL^{d\beta/(\gamma+2\beta)}\right), \end{aligned} \tag{2.94}$$

i.e., the exponent $1/\nu$ in (2.79)–(2.83) is replaced by the exponent $d/(\gamma + 2\beta)$, the inverse of the exponent of the thermal length l_T. Equation (2.94) has been verified for the five-dimensional Ising model [2.76] where $1/\nu = 2$ while $d/(\gamma + 2\beta) = 5/2$.

2.3.7 Finite-Size Scaling for First-Order Phase Transitions

Finally, we turn to finite-size effects at first-order phase transitions [2.97]. In an infinite system, a first-order transition is characterized by a delta function singularity, see Fig. 2.20. For example, if the transition is driven by temperature this singularity is the latent heat; in an Ising magnet, a first-order transition occurs for $T < T_\mathrm{c}$ at $H = 0$ on variation of the field, and hence we get a delta-function in the susceptibility. In finite systems, of course, these delta function singularities are again rounded off [2.98–2.104]. One can understand this behavior most simply by generalizing (2.74) to include the dependence on the magnetic field [2.102]: the weights of the two peaks are no longer equal $\left(\frac{1}{2}\right)$, but rather weighted according to the Zeeman

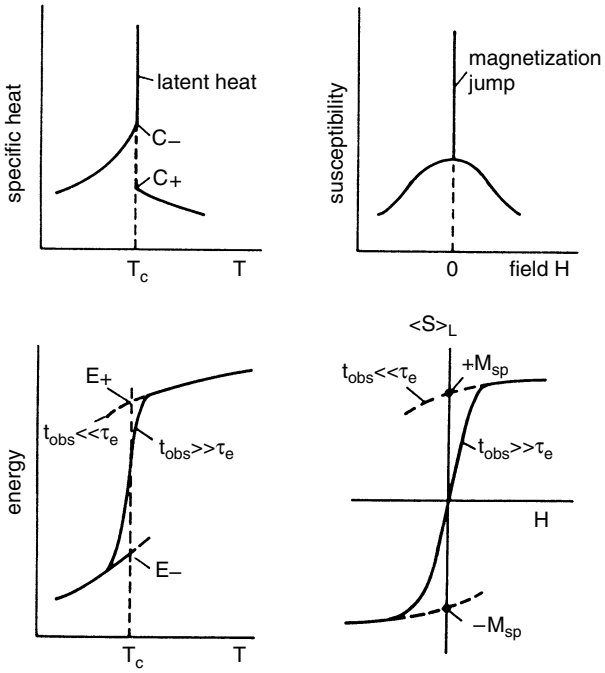


Fig. 2.20 Schematic variation of the specific heat and internal energy with temperature T near a first-order transition at T_c (*left part*). The delta function singularity represents the latent heat $E_+ - E_-$. Also shown is the variation of the susceptibility and the magnetization as a function of field (*right part*). Now the delta function singularity represents the magnetization jump $2M_{sp}$ (*full curves*)

energy Boltzmann factors $\exp(\pm H M_L L^d / k_B T)$. In addition, one must take into account that for $H \neq 0$ the Gaussian peaks no longer occur at $s = \pm M_L$ but rather at $s = \pm M_L + \chi^{(L)} H$. This yields

$$P_L(s) = L^{d/2} \left[2\pi k_B T \chi^{(L)}\right]^{-1/2} \times \left(\frac{\exp\left(\frac{HM_L L^d}{k_B T}\right) \exp\left(-\frac{(s - M_L - \chi^{(L)} H)^2 L^d}{2k_B T \chi^{(L)}}\right)}{\exp\left(\frac{HM_L L^d}{k_B T}\right) + \exp\left(-\frac{HM_L L^d}{k_B T}\right)} + \frac{\exp\left(-\frac{HM_L L^d}{k_B T}\right) \exp\left(-\frac{(s + M_L - \chi^{(L)} H)^2 L^d}{2k_B T \chi^{(L)}}\right)}{\exp\left(\frac{HM_L L^d}{k_B T}\right) + \exp\left(-\frac{HM_L L^d}{k_B T}\right)} \right). \tag{2.95}$$

This approach yields for the magnetization

$$\langle s\rangle_L = \chi^{(L)} H + M_L \tanh\left(\frac{HM_L L^d}{k_B T}\right) \tag{2.96}$$

and the susceptibility

$$\begin{aligned}\chi(H,T,L) &= \left(\frac{\partial\langle s\rangle_L}{\partial H}\right)_T \\ &= \chi^{(L)} + \frac{M_L^2\left(L^d/k_B T\right)}{\cosh^2\left(HM_L L^d/k_B T\right)}.\end{aligned} \tag{2.97}$$

Equation (2.97) shows that the delta function singularity which occurs for $H = 0$ in the limit $L \to \infty$, for finite L is smeared out into a peak of height proportional to L^d and of width ΔH proportional to L^{-d}.

It is important to realize, however, that these considerations apply only if one records the true equilibrium behavior of the system (i.e., applies observation times exceeding the ergodic time). For too short observation time one would observe a single-peak structure for $P_L(s)$ rather than the correct double-peak structure described by (2.95). The ordered state of the Ising ferromagnet then is metastable even in a weak field of direction opposite to the magnetization. The reader who has done Exercise 3.43 of Chap. 3 and who has tried to reproduce Fig. 2.21 will have noticed this already. Thus (2.95)–(2.97) are practically useful only for rather small systems. Figure 2.21 shows that the present description is in quantitative agreement with simulation results for the two-dimensional nearest-neighbor Ising square lattice, when the spontaneous magnetization M_{sp} ($M_L = M_{sp}$ is taken equal to the spontaneous magnetization, independent of L, for $L \gg \xi$) is known exactly [2.105], and no adjustable parameters exist whatsoever.

The field-driven transition in the Ising model is a particularly simple case, since the model possesses a symmetry $\chi(H,T,L) = \chi(-H,T,L)$ even for finite L, and hence the transition is rounded but not shifted by the finite size of the system; it still occurs at $H = 0$. A more general case, which is also more interesting for practical applications, occurs for first-order transitions driven by temperature from an ordered phase at low temperature to a disordered phase at higher temperatures. Obviously, there is no symmetry between the high temperature and low temperature phases, and then we also expect a shift of the effective transition temperature $T_c(L)$ (where the rounded peak of the specific heat representing the smeared delta function of the latent heat has its maximum $C_L^{\max}$) relative to the true transition temperature $T_c(\infty)$,

$$T_c(L) - T_c(\infty) \propto L^{-\lambda}, \quad C_L^{\max} \propto L^{\alpha_m}, \quad \delta T \propto L^{-\theta}, \tag{2.98}$$

where we have defined three exponents λ, a_m, θ for the shift, the height of the peak, and the temperature interval δT over which it is rounded. It turns out, however, that these finite-size effects can again be understood by a simple discussion in terms of

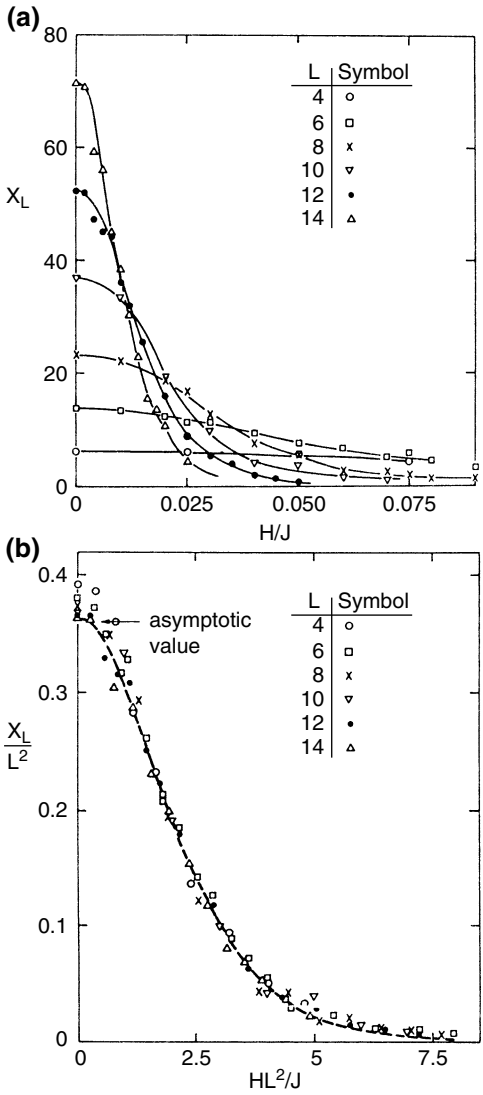


Fig. 2.21 **a** Susceptibility $\chi(H, T, L)$ of nearest-neighbor Ising square lattices at $k_B T/J = 2.1$ plotted versus magnetic field for various L's. **b** Same data replotted in scaled form, $\chi(H, T, L)/L^2$ plotted versus scaled field HL^2/J. *Arrow* indicates the asymptotic value $M_{sp}^2 J/k_B T$ calculated from the exact solution [2.105]. Note that $k_B T_c/J \cong 2.269$ for the Ising model [2.105]. *Broken curve* is the scaling function of (2.97), the additive correction term $\chi^{(L)}$ being omitted. (From [2.102])

thermodynamic fluctuation theory, similar to (2.95)–(2.97). We just have to extend (2.33) to the case where we have a superposition of two Gaussians [$\Delta T = T - T_c$, $T_c \equiv T_c(\infty)$] [2.104]

$$P_L(E) \propto \frac{a_+}{\sqrt{C_+}} \exp\left(-\frac{(E - E_+ - C_+ \Delta T)^2 L^d}{2k_B T^2 C_+}\right) + \frac{a_-}{\sqrt{C_-}} \exp\left(-\frac{(E - E_- - C_- \Delta T)^2 L^d}{2k_B T^2 C_-}\right), \tag{2.99}$$

where the specific heat in the infinite system near T_c behaves as

$$\lim_{T\to T_c^-} C(T) = C_-, \quad \lim_{T\to T_c^+} C(T) = C_+,$$

and the weights a_+, a_- are expressed in terms of the degeneracies q_+, q_- of the two phases and their internal energies E_+, E_- as [2.104]

$$a_+ = q_+ \exp\left(\frac{\Delta T(E_+ - E_-)L^d}{2k_B T T_c}\right),$$
$$a_- = q_- \exp\left(-\frac{\Delta T(E_+ - E_-)L^d}{2k_B T T_c}\right). \tag{2.100}$$

From (2.99) and (2.100) it is straightforward to obtain the energy $\langle E\rangle_L$ and specific heat $C(T, L)$ as

$$\langle E\rangle_L = \frac{a_+E_+ + a_-E_-}{a_+ + a_-} + \Delta T \frac{a_+C_+ + a_-C_-}{a_+ + a_-}, \tag{2.101}$$

$$C(T,L) = \frac{\partial\langle E\rangle_L}{\partial T} = \frac{a_+C_+ + a_-C_-}{a_+ + a_-} + \frac{a_+a_-L^d}{k_B T^2}\frac{\left[(E_+ - E_-) + (C_+ - C_-)\Delta T\right]^2}{(a_+ + a_-)^2}. \tag{2.102}$$

From (2.102) it is obvious that the maximum of the specific heat occurs at

$$\frac{T_c(L) - T_c}{T_c} = \frac{k_B T_c \ln\left[(q_-/q_+)\right]}{E_+ - E_-}\frac{1}{L^d} \tag{2.103}$$

and has a height

$$C_L^{\max} \approx \frac{(E_+ - E_-)^2}{4k_B T_c^2}L^d + \frac{C_+ + C_-}{2}. \tag{2.104}$$

Since the temperature region δT over which rounding of the delta function peak occurs is given just by taking the argument of the exponential functions in (2.100) of order unity, $\delta T \approx 2k_B T_c^2/[(E_+ - E_-)L^d]$, we conclude that the exponents λ, a_m, θ defined in (2.98) are all equal to the dimensionality:

$$\lambda = a_m = \theta = d. \tag{2.105}$$

Thus, since there is no diverging characteristic length to which the linear dimension L could be compared at a first order transition, it is simply the volume L^d that controls the size effects [2.98–2.104]. Figure 2.22 shows results obtained [2.104] for the q-state Potts model [2.106] with $q = 10$, whose Hamiltonian is

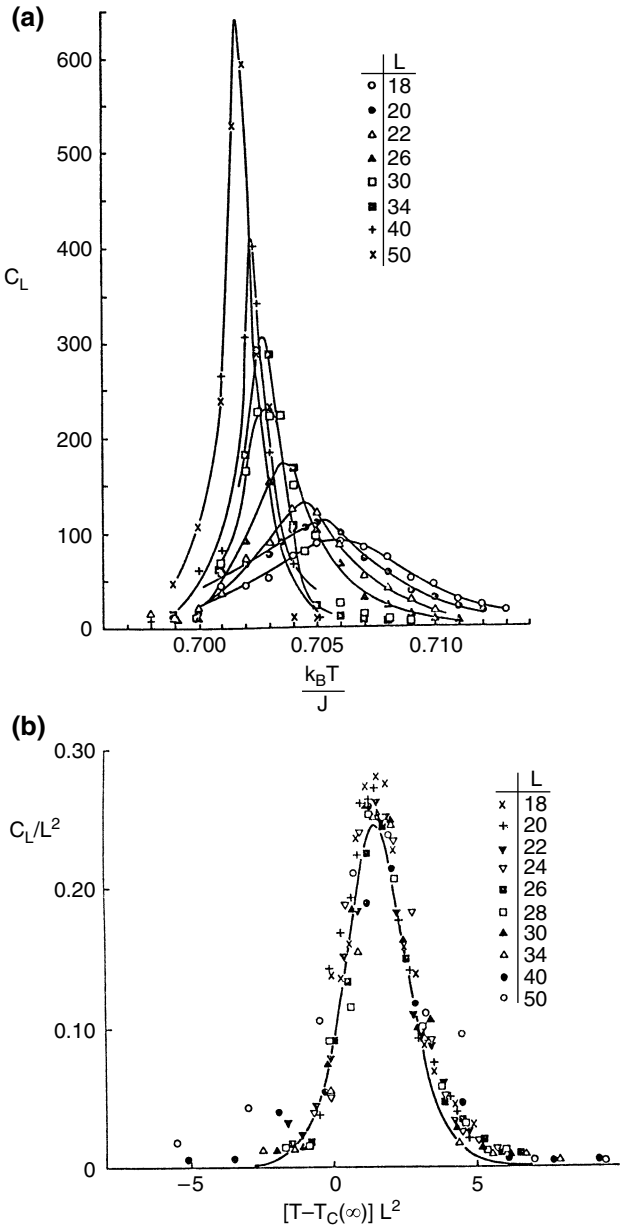


Fig. 2.22 **a** Temperature variation of the specific heat for various lattice sizes, for the ten-state Potts model on the square lattice. Data for some lattice sizes have been omitted in order to preserve the clarity of the figure. **b** Scaling of the specific heat data, plotting C_L/L^2 versus $[T - T_c(\infty)]L^2$. In this model $q_- = q = 10, q_+ = 1$. The *solid curve* results from (2.102), where the additive correction $[a_+C_+ + a_-C_-]/(a_+ + a_-)$ is omitted, and $C_+/C_- = 0.7$ is chosen. (From [2.104])

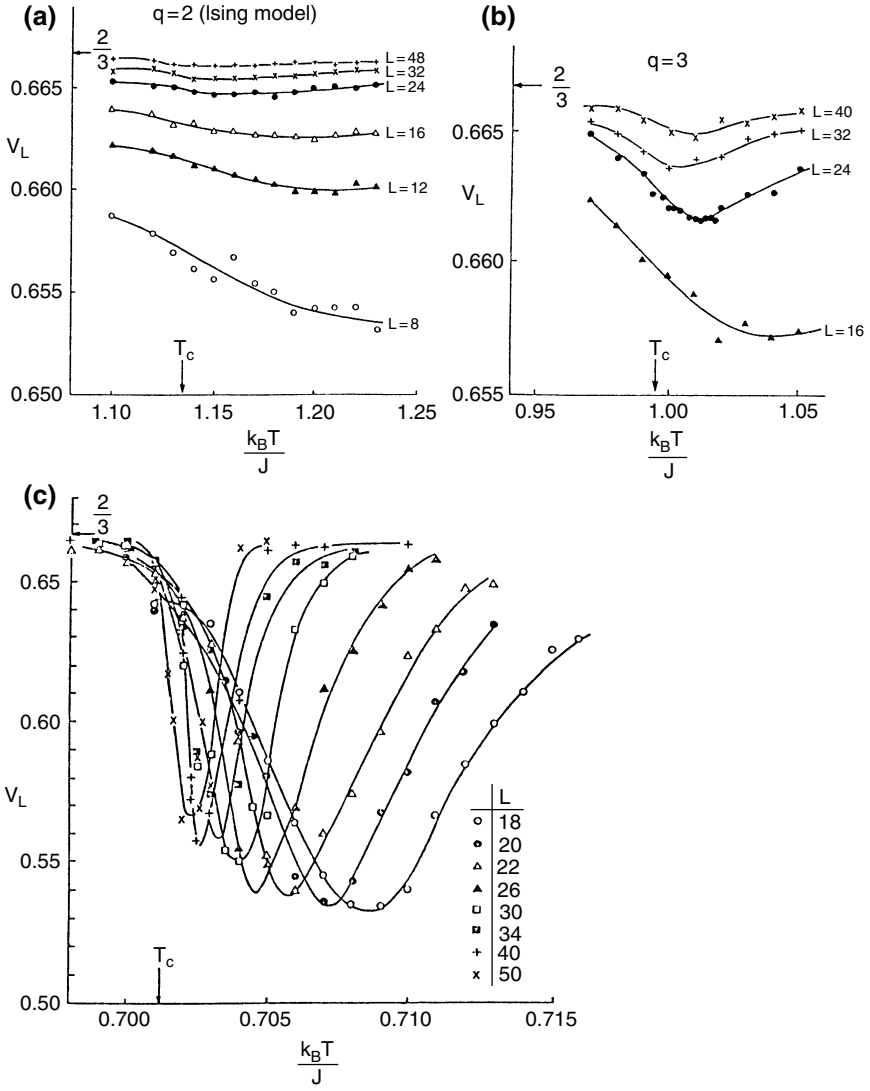


Fig. 2.23 Temperature variation of V_L (2.107) for the q-state Potts model on the square lattice for various choices of L and the cases $q = 2$ (**a**), $q = 3$ (**b**) and $q = 10$ (**c**). Exactly known transition temperatures and limiting values of V_L for $L \to \infty$ are indicated by *arrows*. (From [2.104])

$$\mathcal{H}_{\text{Potts}} = -J \sum_{\langle i,j \rangle} \delta_{S_i S_j}, \quad S_i = 1, 2, \ldots, q. \tag{2.106}$$

For the square lattice E_+, E_- and C_+, C_- can be obtained from exact solutions [2.107] and hence a nontrivial test of (2.99)–(2.105) is possible. It is seen that the phenomenological theory outlined in (2.99)–(2.105) does in fact account for the behavior of the Monte Carlo results nicely. Again some "homework" by the reader (Exercise 3.50) is strongly recommended.

This finite-size scaling theory for first-order transitions also leads to a criterion for systematically distinguishing first-order transitions from second-order transitions. When one computes a quantity V_L defined as [2.104]

$$V_L = 1 - \frac{\langle E^4 \rangle_L}{3\langle E^2 \rangle_L^2}, \tag{2.107}$$

one finds that V_L takes a minimum value $V_L^{\min}$ at the effective transition temperature $T_c(L)$. One can show that for a second-order transition $\lim_{L\to\infty}[\frac{2}{3} - V_L^{\min}] = 0$, even at T_c, while at a first-order transition the same limit measures the latent heat $|E_+ - E_-|$:

$$\lim_{L\to\infty} \left(\frac{2}{3} - V_L^{\min} \right) = \frac{1}{3} \frac{(E_+ - E_-)^2 (E_+ + E_-)^2}{(2E_+ E_-)^2}. \tag{2.108}$$

Figure 2.23 shows [2.104] that indeed the behavior of V_L for $q = 2, 3$ (where the transition is of second order [2.107]) is very different from the behavior of V_L for $q = 10$, where the transition is distinctly first order.

2.3.8 *Finite-Size Behavior of Statistical Errors and the Problem of Self-averaging*

As a last topic connected with size effects we now consider the *finite-size behavior of statistical errors* [2.5]. Suppose we observe a quantity A in n *statistically independent observations*, and calculate its error from [cf. (2.55)]

$$\Delta(n, L) = \sqrt{\left(\langle A^2 \rangle_L - \langle A \rangle_L^2 \right) / n}, \quad n \gg 1. \tag{2.109}$$

We now ask, does this error $\Delta(n, L)$ of the quantity A reduce to zero if we make the simulated system larger and larger, i.e., $L \to \infty$? If this does not happen and Δ reaches an L-independent nonzero limit, we say A exhibits *lack of self-averaging*. If the error Δ reduces to zero, however, we may ask, will the same error be obtained if we simply study larger systems but with less statistically independent observations, such that the total effort in computing time (on a serial computer) is the same. That is, we compare calculations for two sizes L and $L' = bL$, with a scale factor $b > 1$,

and record n observations for the smaller systems and $n' = b^{-d}n$ observations for the larger system. Milchev et al. [2.5] have introduced the notion that a system is *strongly self-averaging* if the error satisfies the relation

$$\Delta(n, L) = \Delta(n', L') = \Delta(b^{-d}n, bL). \tag{2.110}$$

From (2.109) and (2.110) we see that strong self-averaging holds only if the fluctuation satisfies the relation

$$\langle A^2\rangle_L - \langle A\rangle_L^2 \propto L^{-d}. \tag{2.111}$$

If we consider a situation where

$$\langle A^2\rangle_L - \langle A\rangle_L^2 \propto L^{-x_1}, \quad 0 < x_1 < d, \tag{2.112}$$

the quantity A is still self-averaging but not strongly self-averaging [2.5].

Suppose now that the quantity A observed in a simulation is the density of a basic extensive quantity, e.g., energy per site (E) or magnetization per site (M), and we consider for the moment an equilibrium state away from criticality or phase coexistence. Then the distribution of $\delta A = A - \langle A\rangle$ for $L \to \infty$ is Gaussian,

$$P_L(\delta A) = L^{d/2}(2\pi C_A)^{-1/2} \exp\left[-(\delta A)^2 L^d/2C_A\right], \tag{2.113}$$

cf. (2.33) and (2.73), for example. From (2.113) we immediately find that the fluctuation considered in (2.111) is related to the parameter C_A controlling the width of the distribution (if $A = M$ then $C_A = k_B T\chi$, if $A = E$ then $C_A = k_B T^2 C$, etc.):

$$\langle(\delta A)^2\rangle = \langle A^2\rangle_L - \langle A\rangle_L^2 = L^{-d}C_A. \tag{2.114}$$

Thus if $\lim_{L\to\infty} C_A$ exists, A is indeed strongly self-averaging. This property hence holds for quantities such as E, M, etc., for $L \gg \xi$.

The situation differs drastically, however, if we wish to sample quantities which are not spatial averages of a simple density (such as E, M, ...) but quantities which follow from fluctuation relations (such as C, χ, ...), which we obtain from the sampling of E, M using (2.114). We now consider the error of this procedure, again using (2.109) but choosing $(\delta A)^2 L^d$ rather than A as the variable under consideration:

$$\Delta(n, L) = L^d n^{-1/2}\sqrt{\langle(\delta A)^4\rangle_L - \langle(\delta A)^2\rangle_L^2}. \tag{2.115}$$

Since the reduced fourth-order cumulant, defined in analogy to (2.78) as

$$U_L^{\langle A\rangle} \equiv \frac{3\langle(\delta A)^2\rangle_L^2 - \langle(\delta A)^4\rangle_L}{3\langle(\delta A)^2\rangle_L^2} \tag{2.116}$$

vanishes for the Gaussian distribution, we obtain from (2.115)

$$\Delta(n, L) = L^d n^{-1/2} \langle (\delta A)^2 \rangle_L \sqrt{2} = C_A \sqrt{2/n}. \tag{2.117}$$

Consequently, the *relative error* of C_A is independent of L and depends on n only [2.5],

$$\Delta(n, L)/C_A = \sqrt{2/n}. \tag{2.118}$$

Thus, increasing L at fixed n will strongly improve the accuracy of quantities such as E and M, but nothing is gained with respect to the accuracy of χ, C, etc. Thus, it is more economical to choose the smallest size which is still consistent with the condition $L \gg \xi$ and to increase n rather than L to improve the accuracy.

This consideration in turn shows that for large systems it may be better to obtain χ, C, etc. from numerical differentiation rather than from the sampling of fluctuations. For completeness, let us consider this point explicitly. We obtain the susceptibility χ from the field dependence of the magnetization:

$$\chi \cong [M(H + \Delta H) - M(H)]/\Delta H, \tag{2.119}$$

where H has to be chosen such that nonlinear effects are negligible. In order to use (2.119) we must require that the errors of $M(H)$ and of $M(H + \Delta H)$ are much smaller than the difference $\chi \Delta H$:

$$\Delta(n, L) = \sqrt{k_B T \chi / n L^d} \ll \chi \Delta H, \text{ i.e. } 1 \ll n L^d \left(\frac{\Delta H}{k_B T} \right)^2 k_B T \chi. \tag{2.120}$$

The relative error of χ now becomes

$$\frac{\Delta \chi}{\chi} \cong \sqrt{2} \frac{\Delta(n, L)}{\Delta H \chi} = \sqrt{2} \frac{k_B T}{\Delta H} \frac{1}{\sqrt{k_B T \chi n L^d}}; \tag{2.121}$$

the factor $\sqrt{2}$ accounts for the fact that the errors of $M(H)$ and $M(H + \Delta)$ are statistically independent. Since the error is thus proportional to $(nL^d)^{-1/2}$, the susceptibility determined from a derivative of the magnetization is strongly self-averaging. A corresponding consideration holds for the specific heat found as a temperature derivative of the internal energy [2.5].

In the critical region, however, $L \gg \xi$ does not hold; the distributions of the order parameter (magnetization M in an Ising ferromagnet) and energy are no longer Gaussian. Instead descriptions such as (2.79) and (2.80) hold. We now find

$$\Delta M = \sqrt{\left(\langle M^2 \rangle - \langle |M| \rangle^2 \right)/n} = L^{-\beta/\nu} n^{-1/2} \sqrt{\tilde{\chi}(L/\xi)}. \tag{2.122}$$

At T_c, where the correlation length ξ (of the infinite system) is infinite, $\tilde{\chi}(0)$ is a finite constant. Thus, (2.122) is an example of self-averaging of the weak rather than the strong form ($x_1 = \beta/\nu$ is much smaller than d). Using (2.122) to judge the necessary simulation effort one must take into account, of course, that a time of

order L^z (2.86) must elapse between two states in order that they are statistically independent, as shown by (2.87). For the error of the susceptibility χ sampled from magnetization fluctuations, one finds from (2.80) and (2.116) that there is again lack of self-averaging,

$$\frac{\Delta\chi}{\chi} = n^{-1/2}\sqrt{2 - 3U_L(L/\xi)}. \tag{2.123}$$

We thus recognize a structure generalizing (2.118), but now the constant $\sqrt{2}$ is replaced by the smaller quantity $\sqrt{2 - 3U_L(L/\xi)}$.

For the singular part of the energy, $\delta E \equiv E - E(T_c)$, the description similar to (2.79) reads

$$P_L(\delta E) = L^{(1-\alpha)/\nu}\tilde{P}_{(E)}(\delta E L^{(1-\alpha)/\nu}, L/\xi), \tag{2.124}$$

with $\tilde{P}_{(E)}$ a suitable scaling function. From (2.124) one obtains the specific heat as

$$C = L^d\left[\langle(\delta E)^2\rangle_L - \langle\delta E\rangle_L^2\right]/k_B T^2 = L^{\alpha/\nu}\tilde{C}(L/\xi)/k_B T^2. \tag{2.125}$$

For the error of the energy one now obtains again a weak form of self-averaging with $x_1 = (1-\alpha)/\nu$,

$$\Delta(\delta E) = n^{-1/2}L^{-(1-\alpha)/\nu}\sqrt{\tilde{C}(L/\xi)}, \tag{2.126}$$

while for the error of the specific heat a formula analogous to (2.123) applies. As discussed in [2.5], these considerations can be carried over to nonequilibrium Monte Carlo calculations, such as the study of domain growth kinetics [2.46]. We also emphasize that the consideration presented in (2.113)–(2.118) for the errors in importance sampling is essentially the same as that already encountered in our discussion of the simple sampling errors for random walks (2.25)–(2.29).

2.4 Remarks on the Scope of the Theory Chapter

In this chapter we have summarized the main aspects of both "static" (simple sampling) Monte Carlo methods (where different system configurations are generated that are statistically independent of each other) and "dynamic" (importance sampling) Monte Carlo methods. The former have been illustrated with applications to problems such as random and self-avoiding walks, percolation transitions, etc., the latter have been illustrated with applications to lattice problems, with both discrete degrees of freedom (Ising models, Potts models, etc.) and continuous degrees of freedom (XY magnets, the ϕ^4 model, etc.). We have tried to indicate how simulation programs devoted to such applications are organized, and how the resulting Monte Carlo "data" are analyzed. We have discussed in some detail the limitations of these methods due to finite-size effects and due to finite observation time effects,

and have discussed the so-called "statistical errors" from various points of view. We have shown that a detailed analysis of finite-size effects in terms of finite-size scaling theories provides useful information on both second-order and first-order phase transitions. We have also mentioned the application of the Monte Carlo method to studying dynamic phenomena, but have concentrated on static properties here; similarly, the emphasis has been on "bulk properties", and studies of interfaces, surfaces, local properties near defects, etc., have been left aside.

With this restriction in the scope of this book, important problems, such as how one deals with the simulation of quantum-mechanical problems rather than (quasi-)classical Hamiltonians, how one estimates entropy and free energy from importance sampling Monte Carlo, etc., or the analysis of critical and multicritical phenomena with the MCRG method, could not be treated in this chapter and we refer the interested reader to the later chapters in this book, and the references contained therein. We do hope, however, that the present chapter provides a useful guide to the literature for newcomers to the field, enabling them to get started with a few simulations for their own physics problems, and leaving them with the impression that Monte Carlo simulation is a very rich method, which makes interesting contacts with many fundamental aspects of statistical physics, and at the same time is a very versatile and useful tool.

Chapter 3
Guide to Practical Work with the Monte Carlo Method

The guide is structured such that we proceed from the "easy" simulation methods and algorithms to the more sophisticated. For each method the algorithms are presented by the technique of *stepwise refinement*. We first present the idea and the basic outline. From then on we proceed by breaking up the larger logical and algorithmic structures into smaller ones, until we have reached the level of single basic statements. Sometimes we may elect not to go to such a depth and the reader is asked to fill in the gaps.

Since this is a guide to computational physics we feel the need for some general remarks on programming. The technique of stepwise refinement is also useful in actual applications. It introduces a thinking and programming discipline which tends to reduce the number of (logical) errors made. Let us elaborate a little on this. From our own experience we know that nothing is more catastrophic than an untidy program. Bugs may be hidden under untidiness. Though, even in tidy programs there may, unfortunately, be bugs.

Untidy programs may reflect a wild sort of thinking where many little ideas are simply piled up. It is like building a bridge across a river by throwing stones more or less at random into the river. Between each pair of stones we put a statement so that eventually one is able to cross the river. Such a construction is, however, very vulnerable.

Using stepwise refinement for an algorithm we are forced to invent headings for parts of the algorithm and headings for the parts within the parts, as guidance. These are also extremely useful tools for the documentation of the actual program. They come in handy when the program is turned over to a collaborator. If the program is properly documented the colleague will not feel the urge to redo the entire program in order to understand it and trust it.

Interspersed with the text are exercises. These exercises are an integral part of this guide. Without them the whole idea of learning the simulation methods presented in Chap. 2 should be discarded.

K. Binder and D. W. Heermann, *Monte Carlo Simulation in Statistical Physics*,
Graduate Texts in Physics, https://doi.org/10.1007/978-3-030-10758-1_3

There are four different kinds of exercises. Exercises which are marked with ▷, e.g.,

▷ **Exercise 3.1**,

are considered a vital part of the guide. Without doing these exercises an understanding of the material can hardly be reached. They are a "must". Usually they follow immediately after an important step in the development of an argument or an algorithm which has been presented. The reader should not move on until such an exercise has been completed. To mark this even more clearly we have put a

STOP

sign after such an exercise.

The second kind of exercises are those marked with !, e.g.,

! **Exercise 3.2**.

Such exercises are also considered vital to the guide. In these the material presented so far is approached from a slightly different angle. Doing them contributes a great deal to the understanding and the reader is asked to work them out. In most cases they involve variations of an algorithm worked out before either in the text or in one of the ▷ exercises. They usually require some changes to an existing program.

Apart from the regular exercises, which are not marked specifically and are entirely optional, there is a fourth type of exercise marked with a star, e.g.,

⋆ **Exercise 3.24**.

These exercises present problems which are more of a research type and can be safely omitted without losing out on the understanding of the material.

In this guide we will exhibit the algorithms in a pseudo-programming language. This pseudo-programming language is very similar to C, PASCAL, etc. Due to their structure the algorithms which will be discussed present themselves in a more natural and elegant form in this language than in a programming language like FORTRAN. The choice of an artifical language rather than a language in common use also highlights the algorithms themselves: the idea is not buried in a maze of programming statements.

The language is more or less self-explanatory and should be understood by anybody who has experience of at least one computer language. To illustrate the language consider Example 3.1. It simulates throwing a dice. The dice has six faces and we define an integer-valued vector with six components. It is understood that initially the components are zero. We throw the dice N times, as indicated by the "do" loop. For each try we choose at random a number from the set $\{1, 2, 3, 4, 5, 6\}$. In the case that the drawn face is "1" we add a one to the occurrences of face 1, and similarly for the other faces. After N tries have been made we can analyze how often any face appeared in the sample. The statement *analyze-dice-face* stands for a block of statements similar to what is inside the "begin end" for the "do" loop.

Example 3.1

```
integer dice-face(1:6);
do i := 1 to N
   begin
      id := random(1, 2, 3, 4, 5, 6);
      dice-face(id) := dice-face(id) + 1;
   end
analyze-dice-face;
```

It is our experience that most students taking a course in computer simulations know at least the programming language PASCAL. It should therefore be no problem to understand the algorithms and to program them, for example in PASCAL. On most personal computers the language PASCAL is quite efficiently implemented. The use of PASCAL on personal computers has been found to be very successful for teaching purposes. All the algorithms have been tested and run on personal computers. Those who undertake to run the programs on a personal computer will very soon run up against limitations on speed and storage, but for teaching purposes, and sometimes even for research, we find them sufficient in their capabilities.

As already mentioned, the algorithms in this guide are given in a pseudo-programming language. However, we felt that some of the algorithms ought to be given in a "straight" form because while using this guide people sometimes had difficulty in converting the algorithmic idea into a "real" programming language. The conversion is not so trivial at all! It requires a good understanding and working knowledge of the programming language.

In order to overcome this problem, some of the algorithms presented here are given in the Appendix in an executable form. At this point a decision had to be made as to the programming language. As pointed out above, many students start out by learning PASCAL. For teaching purposes this language is most suitable. However, in practice, i.e., when actually doing large scale simulations, in most cases these days the programming language C is used. Therefore some of the algorithms were programmed in C and listings of these are given in the Appendix. For those who prefer Java or Python, etc., it should be no problem to convert the algorithms given in the text into an executable form.

3.1 Aims of the Guide

In Chap. 2 we introduced concepts for the numerical treatment of complex physics problems. In this part we want to fill the abstract concepts with life. We want to understand the concepts more deeply and apply the concepts to problems.

The problems we have chosen for this guide are very elementary. Elementary here does not mean trivial! Indeed, the problems presented here are of a very fundamental nature and are very important for physics. Not just for physics itself, but also for computational physics. They represent a cross section of statistical mechanics problems. Of course, the space in this text is limited and it is impossible to select a broad cross section. The main criterion for the selection of a problem was the didactical value for the application of the numerical concept to the problem.

Basically three approaches to numerically solving a problem in statistical mechanics have been presented:

- Simple sampling
- Biased sampling
- Importance sampling

In a sense, the order in which they are given indicates the increasing complexity of the concepts. They are by no means the only possible concepts. Many other concepts have been developed, such as *umbrella sampling*. Those interested in going beyond the concepts presented here are directed to the general references given at the beginning of the list of references.

The three concepts mentioned above are, however, by far the most prevalent ones in use. Not only that, they are also by far the most pedagogical concepts. Also, since this text is meant as an introduction to Monte Carlo simulation methods, we do not think that more complicated concepts should be elaborated upon here.

In the *simple sampling method* a given problem is approached in the most brutal way. Let us, for the moment, consider an example not from physics. Suppose that you, as a physicist, have been employed by a TV station. Your job is to carry out a poll for the ratings of some shows the station is featuring. Not knowing anything at all about polls you reason that you could hardly ask every viewer who could potentially see the shows. You had a course in statistical physics and know of the power of statistical concepts. You decide to take a sample. You are new in town and the way you go about it is to pick viewers randomly from all over town and question them about how they liked the shows.

In much the same way we could sample the phase space of a problem in statistical mechanics. We take a random sample of points of the phase space and perform averages using the random sample points. Though the concept seems very simple indeed, it is a very powerful one. The whole concept of statistics is lurking behind it. For many problems it even seems to be the only workable approach towards a solution.

In this guide we will treat two fundamental problems in statistical mechanics with the simple sampling method. The first is the *random walk* and other problems of such type. The random walk problem should not be underestimated because of its apparent simplicity. If the reader has really worked through the section on simple sampling where the random walk problem is treated he will have acquired enough experience to master the other methods. The relation of the random walk problem to physics is

rather like that of the fruitfly to genetics. One encounters many difficulties inherent in numerical approaches to the solution of problems in statistical mechanics already in the random walk problem. Understanding the numerical treatment of this problem greatly benefits the general understanding of the working of simulational physics.

Using the basic features of a walk we will branch out to study the *nonreversal* and the *self-avoiding random walks*. These examples will clarify the limits of the simple sampling concept and point the way towards other sampling concepts.

The second problem is that of *percolation*. The application of the concept of simple sampling is rather similar to that of the random walk. The emphasis here is on the analysis. The analysis is twofold. Many problems in physics require the recognition of patterns. One might be interested in the growth of domains, droplets, etc. In the percolation problem the patterns are the clusters. How does one determine the clusters in a given percolation configuration?

The percolation problem will also be used to introduce methods for data analysis. The data analysis we want to emphasize in this text is the determination of:

- The phase transition point
- Scaling

To determine the phase transition point we will use the cumulant method described in Sect. 2.3. This will allow us to determine the phase transition point with very high accuracy by the intersection of the cumulants for various system sizes. The finiteness of the system will also be exploited in the scaling analysis in the form of the finite-size scaling discussed in Sect. 2.3.

An improvement in the level of sophistication is the *biased sampling* concept. Now that you have been in your job for some time you have learned that there are other ways to take a poll. After consulting a street map you found out that there are neighborhoods. Within a neighborhood the habits of the viewers appear to have a certain distribution. The distributions are, however, not all the same. Also, the neighborhoods are not homogeneously distributed on the map of the city. You decide to go along with that and correct for the bias which you have under control.

An example where such a concept can be applied is the sampling of the self-avoiding walk. Applying the simple sampling concept we will find it extremely difficult to sample longer walks. The walks almost always terminate at a very early stage. They terminate because a site to go to next had already been chosen earlier. There is, nevertheless, a way to continue the walk, at the expense of introducing a bias.

After being in the job for more than a while you have smartened up. You decide that even though you have been doing your job well you can reduce the effort you have put in so far. You have learned that often when you picked a household in certain areas they did not have a TV set at all. In some other households the number of hours spent sitting in front of the screen were small. In still others the time spent in front of the screen was very irregular. By now you know where the important contributions to your statistics come from.

Importance sampling applied to statistical mechanics problems is very similar. Though, for most problems, we do not know a priori where the important contributions come from, we can devise an algorithm which surely leads us there.

There is another fruitfly, and that is the *Ising model*. On a lattice with N sites there are 2^N states the system can take on. Most of these states do not contribute much and a simple random sample of the states would yield a very poor convergence. As a matter of fact, it is rather like looking for needles in a haystack, the needles being the states contributing the most to the determination of a quantity.

The treatment of the Ising model with importance sampling will be the major example where the concept is employed. Here we will also pick up again the discussion of data analysis started in connection with the percolation problem.

3.2 Simple Sampling

We begin the guide with examples where the concept of simple sampling is applied. Two major examples will be used to exhibit the application of the concept to the numerical solution of the random walk and the percolation problem. We shall also branch out to treat the nonreversal random walk and the self-avoiding random walk.

3.2.1 Random Walk

In the first example we want to treat the random walk (RW) problem with the simple sampling Monte Carlo method. What we will be after, in particular, is to calculate the average *end-to-end distance* $\langle R \rangle$ of a random walk as a function of the number of steps.

An algorithm to calculate the end-to-end distance and other quantities of the random walk problem requires several input parameters. First of all one needs, of course, the step length N of the walker. This parameter we want to vary to establish the relation between the step length and the average end-to-end distance. As a second parameter we need to specify how many samples of N-step random walks we want to generate. We shall denote this by *n-of-samples*.

In a block notation the Monte Carlo simple sampling of random walks then looks like Algorithm 3.1.

This is the basic building block for the simple sampling of the end-to-end distance. The outer loop controls how many times we want a sample of the end-to-end-distance of a walk of N steps. The generated result of one trial is accumulated and later on an average is taken over the entire sample of size *n-of-samples*. No preference is given to any of the trials. All trials of the sample carry exactly the same weight. The average we take is then the simple arithmetic average. Let r_i denote the end-to-end distance of trial i, then

Algorithm 3.1 Random walk

```
do sample := 1 to n-of-samples
  begin
    do step := 1 to N
      begin
        generate-one-step
      end
  end
  accumulate-results
end
```

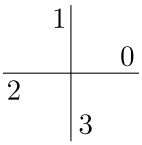
$$\langle R \rangle = \frac{1}{\textit{n-of-samples}} \sum_{i=1}^{\textit{n-of-samples}} r_i . \tag{3.1}$$

That the end-to-end distance is calculated by an arithmetic average is not the real signature of the simple sampling for the end-to-end distance of random walks. The real signature is the generation of a trial walk.

The innermost loop in Algorithm 3.1 is the loop which actually creates one realization of a random walk with a step length of N. For definiteness, we study the random walk in two dimensions on the simple square lattice (Fig. 2.1). Every random walker will start from the same origin $(x, y) = (0, 0)$. This is just for convenience, because it makes the calculation of the end-to-end distance easier.

The task at hand is to resolve the heading *generate-one-step*. Remember, the concept of simple sampling implies that on no level is any preference given whatsoever.

At each step a random walker has the choice to go in one of four directions with a probability of $1/4$. Each direction is *equally* probable. Let us label the directions as



We realize this algorithmically by generating a random number which can take on the values 0, 1, 2 or 3 with equal probability

$$ir := \textit{random}(0, 1, 2, 3).$$

Once we have chosen a direction we can advance the random walker to that position, assuming that he was at position (x, y) before:

```
case ir
  0 :  x := x + 1.0;
  1 :  y := y + 1.0;
  2 :  x := x - 1.0;
  3 :  y := y - 1.0;
end
```

Algorithm 3.2 Random walk

```
do sample := 1 to n-of-samples
  begin
    x := 0; y := 0;
      do step := 1 to N
        begin
          ir:=random(iseed)*4;
          case ir
          0: x := x + 1.0;
          1: y := y + 1.0;
          2: x := x − 1.0;
          3: y := y − 1.0;
        end
      end
    accumulate-results
  end
```

We have almost resolved the block labeled *generate-one-step*, i.e., the generation of one basic step of the random walker. What is left is the question of how we get the random numbers. For this we assume for the moment that there is a procedure which supplies random numbers distributed uniformly between zero and one. We can get the desired four numbers by using the rounding and truncating features of the programming language. If we assign a real number to an integer variable the real number will be either truncated to the integer part by discarding the fractional part or rounded to the nearest integer. What we need at the moment is the truncation so that the number 3.99 . . . is truncated to 3. In our artificial language we assume that every time a real-valued variable is assigned to an integer-valued variable the fractional part is lost. Otherwise we have to specify explicitly that the variable is to be rounded to the nearest integer. Now we can write

$$ir := random(iseed) * 4.$$

Here *iseed* is a seed which is needed to start up the random number generator. After all the discussion the algorithm to perform a simple sampling of end-to-end distance random walks is Algorithm 3.2.

For those who have not yet worked with random number generators, but also for those who have (!), we shall supply a generator. As an example of a random number generator we take the function in Algorithm 3.3.

Some comments are due at this point. The reader is asked not to manipulate and use other values for the parameters *mult* and *modulo*. The routine should be started with an odd-valued seed. The last comment concerns the modulo function *mod*, which we assume to exist.

Algorithm 3.3 Modulo random number generator

```
real function random(ibm)
  integer mult, modulo, ibm;
  real rmodulo;
  begin
    mult := 1277;
    modulo := 2 ↑ 17;
    rmodulo := modulo;
    ibm := ibm ∗mult;
    ibm := mod(ibm, modulo);
    random := ibm/rmodulo
  end
```

▷ **Exercise 3.1**. Program Algorithm 3.3 for the simple sampling of random walks. Vary the step length and the sample size. Record the average end-to-end distance in the x and y directions (separately!), the fluctuation of these quantities and the average execution time of the program as a function of the step length.

STOP

Before continuing the reader is asked to work out the above exercise. We feel that the exercises marked with ▷ are vital to the understanding of the material presented in this text. They should not be skipped.

Let us examine the results on the end-to-end distance of random walks generated with the simple sampling outlined above. The x and y directions of the walker are equally probable. We should get the same result for the average end-to-end distance in the x direction as for y direction. Figure 3.1 shows results obtained using the random number generator given above. The end-to-end distances are almost the same. We also *know* that the result must be

$$\langle R^2 \rangle \propto t, \tag{3.2}$$

but from the numerical treatment we have to conclude a different law.

What went wrong? The concept itself cannot be wrong! At this point we need to concern ourselves again with the generation of random numbers. The random numbers are at the very heart of the Monte Carlo method, whether it be based on simple, biased or importance sampling. The success, i.e., the correctness of the results, of a Monte Carlo simulation hinges crucially on the quality of the generated random numbers.

Let us look at an even more sensitive measure. We know from Sect. 2.3.8 that the end-to-end distance exhibits a lack of self-averaging. The error in the end-to-

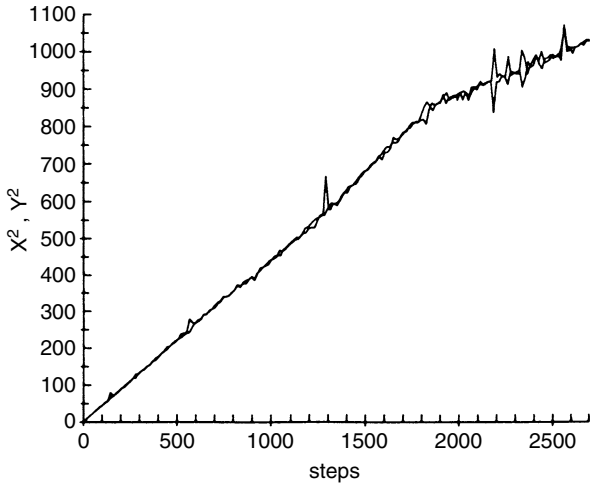


Fig. 3.1 Average end-to-end distance for the random walker. Plotted separately are the x and y directions. The results were obtained using the random number generator of Algorithm 3.3

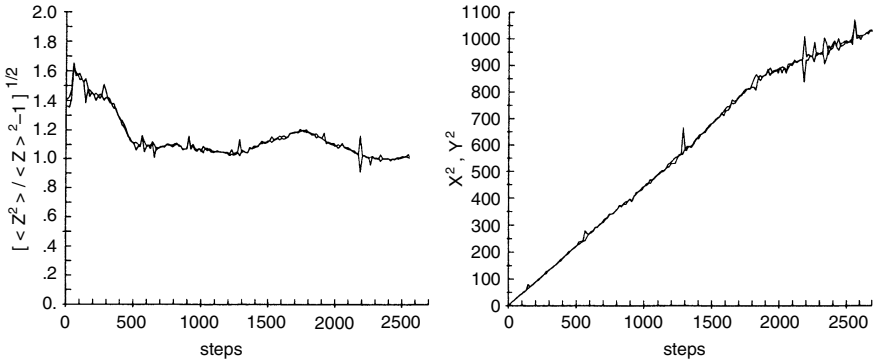


Fig. 3.2 *Left*: The error of the end-to-end distance, which should be constant, as a function of the step length. Here z stands for either x^2 or y^2, respectively. *Right*: The corresponding average end-to-end distances. The results were obtained using the random number generator of Algorithm 3.3

end distance must be independent of the step length. The result on the error of the end-to-end distances is shown in Fig. 3.2. The error is not constant!

We do not want to go into great detail. There are other texts dealing with the problem of the generation [3.1–3.5] of random numbers, or more precisely pseudo-random numbers. Here we just present a brief introduction to the subject.

The most popular and prevalent generators in use today are the *modulo generators*. An example of this type of generator was given above. Even though we have had a bad experience with this generator we should not despise it altogether.

The modulo generator is attributed to *Lehmer* [3.6]. Let m, a, c, x_0 be integers with $m > x_0$ and $a, c, x_0 > 0$. The basic idea used to generate a random number is a folding mechanism. This is very much like the mappings one studies for chaotic properties. The constructed sequences of numbers appear to be highly unpredictable.

As the name suggests, the key to the folding is the modulo function. The pseudo-random number x_i of the sequence $(x_i)_{i=1,\ldots}$ is obtained from x_{i-1} by the recursion relation

$$x_i = ax_{i-1} + c \pmod{m}.$$

It can be guessed from our first choice of parameters m, a and c that the statistical properties of the generated numbers depend crucially on them. A bad choice ruins the results of a Monte Carlo simulation. Several choices for the parameters have been proposed [3.1–3.5], and the reader is asked to consult the cited texts.

One remark, though, should be made here. Numerical treatments of problems often require a great many random numbers. Hence the period after which the sequence repeats must be quite large. It must be very much larger than the number of random numbers required in the simulation, otherwise erroneous results appear. The reader can easily convince himself of that by doing the random walk simulation with parameters chosen such that the period is small. The data shown in Fig. 3.1 show that we eventually reached the cycle length of the random number generator.

Most computer installations and some programming environments provide a random number generator. From the experience above it should be clear that you should not trust a generator before it has been tested by you! Unfortunately most statistical tests fail to show correlations which appear later in a bias of the results of simulations of problems in statistical mechanics. Probably the best test is to apply the generator to a real problem where the results are known or where it can be inferred that something must be wrong with the generator [3.7].

In our example (Figs. 3.1 and 3.2) we see that the correlations in the random numbers can sometimes show up in a subtle way. Some quantities may be affected only very weakly. We see that the average end-to-end distance comes out as expected for the walks of smaller lengths where no problem with the cycle length existed. The error in the end-to-end distance tells a different story and is a more sensitive measure in these circumstances.

There is a quick test which no generator should fail. The quality of a random number generator can be seen if one tries to completely fill a d-dimensional lattice. We have seen above that we can map the interval $(0, 1)$ to the discrete set $\{0, 1, \ldots, L-1\}$ of integer numbers. This mapping allows us to construct random vectors $(x_1, \ldots, x_d)$ of integers in a d-dimensional hypercube. Each site reached with a random vector derived from the generator is marked. A good generator must cover the entire hypercube.

The result of such a test in two dimensions using the modulo generator is shown in Fig. 3.3. Clearly the generator performs very badly, i.e., shows duplet correlations. Only parts of the lattice are reached. Many sites are not visited at all! In three dimensions the situation is even more dramatic.

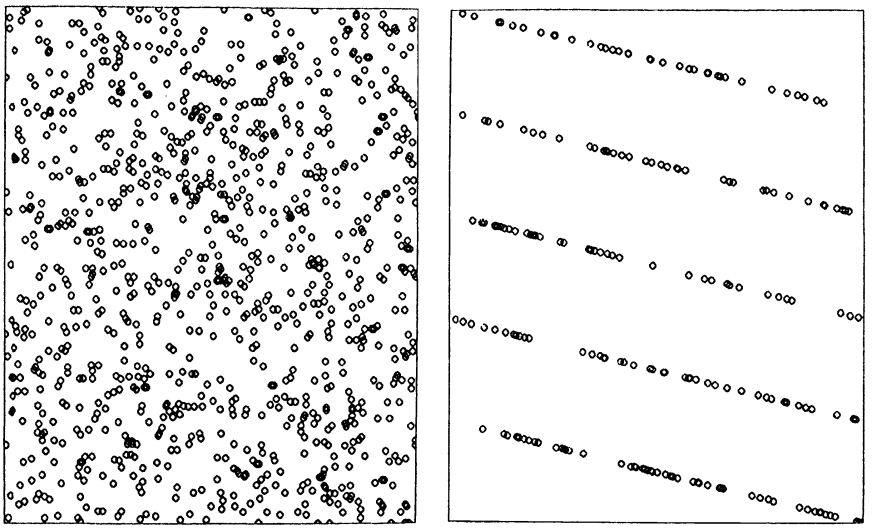


Fig. 3.3 Example of a test of a random number generator. The *circles* are the sites of a two-dimensional lattice which were reached by vectors obtained from the random number generator. The rhs figure shows strong correlations among the (pseudo-)random numbers

! **Exercise 3.2**. Write a program which performs the above test for a random number generator. How should the number of unvisited sites decay with the number of attempts to fill the lattice?

It should be noted that the described method is actually used in Monte Carlo algorithms. Often it is required to go randomly through a lattice of spins. The results show that when using a bad random number generator certain lattice sites would never be visited and the results are therefore biased.

! **Exercise 3.3**. Verify numerically the Einstein law

$$\langle R^2 \rangle \propto t$$

by performing the simulation in two, three and four dimensions.

! **Exercise 3.4**. *Biased random walk*. Algorithm 3.1 can easily be modified for the study of the biased random walk. To make the matter somewhat easier we work in one dimension. In this problem we choose the $+1$ steps with a probability p and the -1 steps with a probability q such that $p + q = 1$. After having modified Algorithm 3.1 for this problem, are we still doing a simple sampling?

Exercise 3.5 Use the algorithm given in Example 3.1 to test the random number generator given in the text. Increase the number of bins and test for uniformity and skew.

Exercise 3.6 *Diffusion limited aggregation*. In the diffusion limited aggregation problem one starts with a seed on a d-dimensional lattice. A random walker is started

from the edge of the lattice. If the random walker gets to one of the nearest-neighbor sites of the seed it sticks and the two form an aggregate of two sites. A new random walker is released and sticks to the aggregate if it gets to one of the nearest-neighbor sites, etc. Write a program to generate a diffusion limited aggregation structure.

Exercise 3.7 Construct an algorithm to generate random walks of length N which does not utilize the "case" statement.

Exercise 3.8 Determine the number of loops a random walk makes as a function of the step length N.

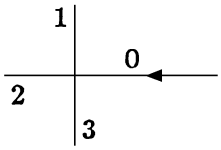
Exercise 3.9 Invent an algorithm for the random walk problem which runs on a parallel machine with p processors which you are allowed to connect whichever way you like.

3.2.2 Nonreversal Random Walk

Nonreversal walks are not allowed to immediately fold back. Each step has to be a "forward" step. By "forward" step we mean that the walker is not allowed to make an immediate reversal. It is, however, allowed to cross its own path or at some later time retrace some of the steps made earlier (Fig. 2.1).

For the numerical treatment of the nonreversal random walk (NRRW) by simple sampling we can follow at least two possible routes. Overall the algorithm is the same as that for the simple sampling of random walks. Only the generation of a walk, i.e., one try, has to be modified to incorporate the constraint.

One possibility is to ignore at each step that the walker came from somewhere (except at the origin). If the step that the walker takes goes back to the one it came from, another try is made until a "forward" step is found. Here we choose at each step one of the four sites in our two-dimensional problem with probability $1/4$:



Suppose that by chance we choose the direction labeled "0". An immediate reversal to the site visited before follows. As a result, a new trial for a step is made until a valid direction is picked. Here, in the example, as soon as one of the numbers $\{1, 2, 3\}$ appears a valid step is made.

The other possibility is to remember where we came from and choose only from the three possible neighboring sites. Here the probability is taken to be $1/3$ for each of the three possible directions. There are always three directions even if all the sites the walker can go to have been visited before at some time. Only the one visited *immediately* before the current step is prohibited.

Algorithm 4 Nonreversal random walk

```
x := 0; y := 0;
x_pre := 0; y_pre := 0;
do step := 1 to N
  begin
    repeat
      x_temp := x; y_temp := y
      generate-one-step (x, y)
    until ((x_pre ≠ x) or (y_pre ≠ y));
    x_pre := x_temp
    y_pre := y_temp
  end
```

This route requires the identification of the possible nearest-neighbors and the creation of a list of these. From this list with three entries one member is chosen with equal probability.

For now we take the ignorant approach and may knock at closed doors several times. This approach requires little modification to the basic random walk algorithm given in the previous section. We could, for example, introduce new variables

$$x_{\text{pre}},\ y_{\text{pre}}$$

to keep track of the preceding position and to test whether the chosen new position is the same as the preceding one. If so, we repeat the generation of the step until the preceding and new positions are not identical. See Algorithm 3.4.

▹ **Exercise 3.10**. Program Algorithm 3.4 for the sampling of nonreversal random walks. Vary the step length and the sample size. Record the average end-to-end distance in the x and y directions (separately!), the fluctuations of these quantities, and the average execution time of the program as a function of the step length. Compare your numerical results with the exact answers.

STOP

! **Exercise 3.11**. Develop an algorithm for the nonreversal random walk problem which chooses at each step always one of the $2d - 1$ possible nearest-neighbor sites on a simple d-dimensional lattice.

Exercise 3.12 Compare the performance of Algorithm 3.4 and the algorithm from Exercise 3.11.

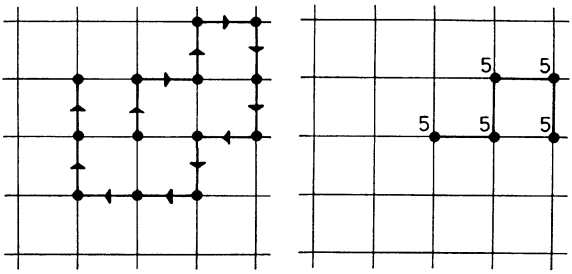


Fig. 3.4 *Left*: An example of a self-avoiding random walk. *Right*: A random walker carrying a label of 5 which he leaves at all places he visits

3.2.3 Self-avoiding Random Walk

The generation of self-avoiding random walks requires a little more algorithmic complexity than the random or nonreversal random walks. In the random walk problem no constraint was placed on the possible sites the walker could visit. At each step the walker has completely forgotten where it came from. Each site is considered an origin from which there are $2d$ (for a simple d-dimensional lattice) possible directions to go in. The self-avoiding random walk (SAW) is not allowed to go to all sites. It is not allowed to cross its own path and the walker is not allowed to make an immediate reversal. At each site it *locally* remembers where it came from and has only a maximum of $2d - 1$ allowed choices. One choice of the $2d$ nearest neighbors is always lost due to the nonreversibility constraint. The local environment may prohibit some of the other choices due to the condition that the path must not cross itself. There are situations where there are no choices left at all!

An example of a walk is given in Fig. 3.4 (cf. also Fig. 2.1). Starting from the origin, the random walker performed 12 steps. If the 13th step were to go to the right or down then the walker would go to a previously visited site. Both cases are not allowed. In the second case it would be an immediate reversal. In the first case the walker would visit a site twice and as a consequence the walk *terminates* with a length of 12 steps.

That the walk terminates once a site out of the $2d - 1$ neighboring sites is chosen which has been visited before is the key to the simple sampling of self-avoiding walks. We could also argue that the site was forbidden anyway. So why not choose out of the ones we can actually go to? We will follow up on this idea in the next section where we discuss the application of biased sampling to the self-avoiding walk problem.

In the simulation of random walks we did not take any notice of the actual path or trace of the walk. For the self-avoiding random walk we have to check at each step whether it is a valid step. We have to record the trace of the walker on the lattice (Algorithm 3.5). The generation of the walk itself is the same as for the random walk problem.

The only new part which has come in is that we do not generate exactly N steps, as for the random walk problem, but an unpredictable number of steps. The walk can terminate at any step due to the constraints. To provide for an unpredictable number

Algorithm 3.5 Self-avoiding random walk

```
do sample := 1 to n-of-samples
  begin
    step := 0;
    repeat
      generate-one-step;
      step := step + 1;
    until (step-invalid or step = N)
  accumulate-results
end
```

of steps we have replaced the unconditional loop over N steps with a loop where after the generation of a step it is checked whether an invalid move has been made or the maximum number of steps N has been reached. In both cases the walk terminates. The length of the walk in steps is always less than or equal to N. For the random walk it was always equal to N.

How do we check that the step made is a valid one? Let us assign to each walker a unique label. To the first walker we assign the label "1", to the second, the label "2", etc. Such a label comes automatically with the enumeration of the walkers by the outer loop. As the walker advances it labels the sites it has visited with its unique label (Fig. 3.4). On advancing to the next site it checks whether the site carries the same label as its own. If so, then this site has been visited before and the self-avoiding condition applies with the termination of this particular trial of a walk.

We are here at a branching point. We have to decide how to generate a single step. The choices we have are the same as for the generation of the nonreversal random walk. We can either be ignorant and use the algorithm for the generation of a single step as for the random walk problem or we can a priori choose only out of the $2d - 1$ potentially valid sites. For the moment let us be ignorant.

The labeling scheme does not yet allow us to distinguish whether the site was the one previously visited or an older one. If it is the previously visited one then we have to try again. In the other case the walk terminates.

The labeling scheme does, however, allow us to determine whether the step is valid or not. We may take advantage again of the uniqueness of the label. Before advancing to the next site we step up the label of that site by one. Once the walker has advanced it can check whether it advanced to a site carrying the label of the present walk (in this case the walk terminates) or the label plus one (in this case it has reversed its step and another try is made).

One way to implement the labeling scheme for the trace of the walk is to introduce an array. For the other examples we did not require such an array, neither for the random nor for the nonreversal random walk. What are the dimensions of the lattice? At most the walker can travel a distance N. Even though to travel a distance N is highly unlikely, we have to provide for the extremum. Since the walker can go in any direction we dimension the lattice as (Algorithm 3.6)

Algorithm 3.6 Self-avoiding random walk

```
integer lattice(−N : N, −N : N);
do sample := 1 to n-of-samples
  begin
    step := 0;
    x := 0; y := 0;
    x_c := 0; y_c := 0;
    repeat
      repeat
        generate-one-step(x_new, y_new);
      until lattice(x_new, y_new) ≠ sample + 1;
      if lattice(x_new, y_new) = sample then
        terminate := true
      else
        begin
          lattice(x, y) := sample;
          x := x_c; y := y_c;
          lattice(x, y) := sample + 1;
          x_c := x_new; y_c := y_new;
          step := step + 1;
        end
    until (terminate or step = N);
    accumulate-results
  end
```

integer *lattice*$(-N : N, -N : N)$

▹ **Exercise 3.13**. Program Algorithm 3.6 for the simple sampling of self-avoiding random walks. Print out the lengths (in number of steps) where the walks terminated. Plot the number of self-avoiding walks as a function of the number of steps N. Obtain also the entropy S by the fraction of successful attempts $W(N)$ using the relation $S - S_0 = \ln[W(N)]$ where S_0 is the result for the random walk.

STOP

We do not know when a walk terminates. Different trials will result in different lengths of a walk before it terminates. From the above exercise we have gained some experience of how much the length varies. If we want to obtain results on the end-to-end distance for a length of N (N large) there are many walks terminating with a length less than N, see (2.16). For large N it becomes nearly impossible to compile enough statistics! Instead of discarding trials less than the desired length, we record them and get results for the end-to-end distance for walks of lengths between 1 and N. The statistics will, however, be different for the different lengths. The short lengths will have by far larger sample sizes than the longer lengths. This makes the simple sampling of the self-avoiding walk rather impossible for lengths of, say, 100 and larger.

▹ **Exercise 3.14**. Program Algorithm 3.6 for the sampling of self-avoiding random walks. Instead of printing out the lengths where the walks terminated record the end-to-end distances in the x and y directions separately. Compute also the fluctuations of the end-to-end distances and record the execution time necessary to gain sufficiently good statistics for the largest step length you can manage on your computer system. Check that the error is under control.

STOP

! **Exercise 3.15**. Develop an algorithm for the SAW problem which chooses at each step always one of the $2d - 1$ possible nearest-neighbor sites on a simple d-dimensional lattice.

Exercise 3.16 Can you think of another way to keep track of the path of the walker. (Hint: use a stack.)

Variations on a theme

Exercise 3.17 *Reptation*. The connection between polymers and self-avoiding walks leads on to another idea for generating self-avoiding walks. Assume we have a self-avoiding configuration, or conformation in polymer language, of length N steps. Remove a "link" at one end of the chain. At the other end of the chain pick a random direction (out of the $2d - 1$ potential ones) and try to paste there the removed link in this direction. If the self-avoiding condition is violated do not accept this move. Otherwise a new configuration is found and the procedure is repeated. What sampling concept applies to this algorithm? To be sure, wait for the answer to this question until the end of the guide.

3.2.4 Percolation

The simple sampling method for the percolation problem appears, on the face of it, to be an even simpler problem than the random walk problems discussed in the previous sections. Simple sampling here means the generation of configurations and the analysis of these with a uniform statistical weight attached to them. Each configuration is treated on an equal footing.

Let us briefly recall that we are interested in finding the point of the geometrical phase transition p_c for the percolation problem (Sect. 2.3.1). Below p_c there are only finite clusters. Above p_c there is at least one infinite cluster. This geometrical phase transition is the analog to the second-order phase transition at T_c for thermal systems. We shall take the opportunity here and begin with the study of phase transitions and their analysis apart from the problems specific to percolation theory.

Algorithm 3.7 Percolation

```
do no-of-config := 1 to N
  begin
    generate-a-configuration;
    analyze-the-configuration;
    calculate-quantities;
    accumulate-results;
  end
perform-averaging;
```

Algorithm 3.8 *generate-a-configuration*

```
do i := 1 to L
  do j := 1 to L
    if p < random(iseed) then
      lattice(i, j) := 0
    else
      lattice(i, j) := 1;
```

Before addressing the issue of exponents we start with the determination of the threshold p_c. For the simple sampling we want to generate N realizations of a percolation configuration and average over quantities derived from these configurations. The basic algorithm for doing so is given as Algorithm 3.7.

In the percolation problem, more precisely the *site percolation problem*, we consider lattices (for convenience we again assume a two-dimensional lattice) filled with a probability p. One filling is considered a configuration in the simple sampling scheme. To generate one configuration we have to visit each site of the lattice once and decide whether we fill it or leave it unoccupied. This can be done as follows

Here *random* is a real function which returns a random number in the interval (0, 1) and *iseed* an integer which, upon first calling, initializes the random number generator.

▹ **Exercise 3.18** Program Algorithm 3.8 for the generation of one configuration of the site percolation problem. Print out the configurations for various choices of the probability p with which the lattice was filled and inspect them visually.

STOP

The analysis of a generated configuration is usually done by counting the number and size of clusters in the configuration. As described earlier, one defines the order parameter of the phase transition using the clusters.

We now need to develop an algorithm to identify a cluster in a generated configuration of sites. Let us recall that a *cluster* is defined as follows: Two occupied sites belong to same cluster if they are nearest neighbors. Hence, a cluster is a subset

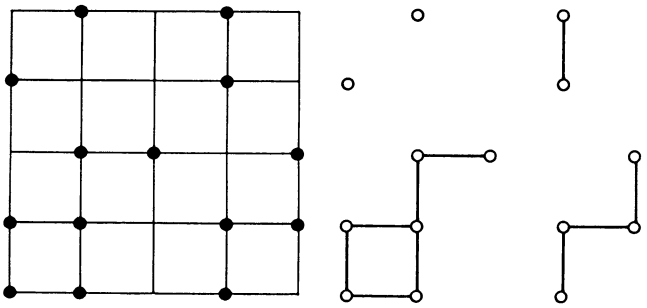


Fig. 3.5 *Left*: A site percolation configuration. *Right*: The subgraphs of the above configuration

of the occupied sites such that for each site belonging to the subset we can find a nearest-neighbor site belonging to that subset.

We may view the occupied sites of the lattice as a graph. This graph can have many disjoint subgraphs, as illustrated in Fig. 3.5. There are graphs containing only a single site, clusters containing two sites, etc. The task at hand is to enumerate all subgraphs, i.e., how many subgraphs are there with how many sites?

Before attempting to identify and enumerate all the clusters in a given configuration we set ourselves a somewhat easier task: identify an infinite cluster. Even simpler! Not the entire cluster but only a part. In other words, is there a path which leads us from one side of the lattice to the opposite side. The probability of finding such a path defines also an order parameter P_s, which can be used to determine the transition point. Below p_c the probability of finding such a path is zero; only finite clusters exist, which cannot lead, in the thermodynamic limit, to the opposite side. Above the threshold an infinite cluster exists which allows one to cross from one side to the opposite side, and the order parameter is one. At p_c the order parameter jumps from zero to one.

▷ **Exercise 3.19**. Invent an algorithm which decides whether in a given configuration of the site percolation problem one can cross from one side of the lattice to the opposite side of the lattice. Such a cluster is called the spanning cluster. Determine the order parameter P_s for various lattice sizes and probabilities p.

STOP

Those who did Exercise 3.19 will have found that the order parameter has a smooth behavior for the lattice sizes investigated. There is no jump at a unique probability p from zero to one. We have encountered for the first time *finite-size effects* in a phase transition problem. Let us postpone the discussion of finite-size effects a little bit and return to the question of the identification of clusters in a given configuration. From the experience with the order parameter P_s, we expect to find a similar feature for the order parameter P_∞ derived from the largest cluster.

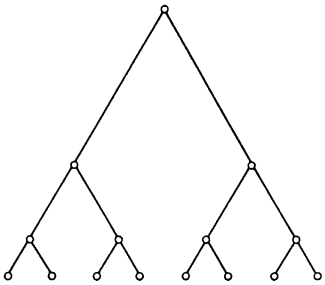


Fig. 3.6 A binary tree on which we study percolation configurations

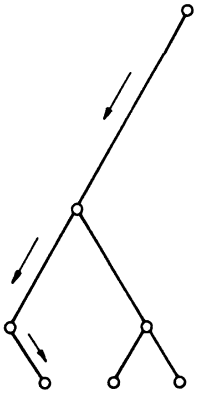


Fig. 3.7 One cluster of a percolation configuration on a tree. The path to find the first leaf or dangling end of the cluster is indicated

Probably the most natural way to identify all the clusters is a *recursion algorithm*. To illustrate the idea we simplify the problem. Instead of working on a square lattice we work on a binary tree. An example of such a tree is displayed in Fig. 3.6.

For each site there are only two nearest neighbors in the downward direction and one predecessor in the upward direction. The percolation problem itself remains defined as above. Each site is visited and is set occupied, with a probability p, or left unoccupied with a probability $1 - p$.

Consider a configuration of occupied and empty sites on the tree. Pick one occupied site. To determine which of the sites belong to the same cluster we traverse the tree until all possible paths have been explored. Since each site has at most two occupied nearest neighbors in the downward direction, there are at most two possible directions to be explored. We shall call these directions "left" and "right".

The idea used to visit all the sites belonging to a cluster is to first travel along the cluster as far as possible, i.e., until a leaf of the tree or, in the terminology of the percolation problem, a *dangling end* has been found. The strategy for finding such a dangling end is to go to the left as far as possible. In Fig. 3.7 we show such a path. Once we have reached a point where we can proceed no further to the left, we go to the right one step, if possible, and then continue to go left, etc. Eventually we will reach a site which has neither left nor right neighbor. This site is a dangling end.

Algorithm 3.9 *analyze-the-configuration* for a binary tree

```
burned := 0;
repeat
  select-an-unburned-site(i)
  n := 1
  go-through-tree(i, n)
  burned := burned + n
until occupied-sites = burned
procedure go-through-tree(i, n)
begin
  if tree(i, 1) ≠ 0 then go-through-tree(tree(i, 1), n);
  if tree(i, 2) ≠ 0 then go-through-tree(tree(i, 2), n);
  tree(i, 1) := 0;
  tree(i, 2) := 0;
  n := n + 1;
end
```

We remove this site, raise the counter for the number of sites in the cluster by one and go back to the site we came from.

Now we try to go up by going to the right. If this is not possible then the removal of the dangling end has made the current site a dangling end, which can be removed. The algorithm proceeds to work upwards on the tree until we are back at the site we started out from, and hence all sites belonging to the cluster have been visited.

The algorithm produces dangling ends which can be stripped or *burned* from the cluster and eventually the entire cluster is burned away. Only those sites remain which are not connected to the current cluster. A fresh site can be selected and the procedure of burning sites is repeated. The algorithm stops if there are no sites left and the whole tree is consumed.

In Algorithm 3.9 we have organized the storage of the tree with N sites in a two-dimensional array tree $(1 : N, 1 : 2)$. The first index gives the site and the second index gives the left or right neighbor. The value of the array elements can be either 0 or an integer number out of the set $\{1, \ldots, N\}$. If we find a 0 then the site is not connected, otherwise we have a pointer to the next node or site.

▹ **Exercise 3.20**. Program Algorithm 3.9 for the identification of clusters of the site percolation problem on a binary tree.

STOP

That we elected to exhibit the burning algorithm for the percolation problem on a binary tree was, of course, intentional. On a regular lattice the algorithm becomes slightly more complicated. The general idea for traversing a cluster remains the same. The simplifying feature of the tree is that it has no loops! On a regular lattice we can have loops as shown in Fig. 3.8.

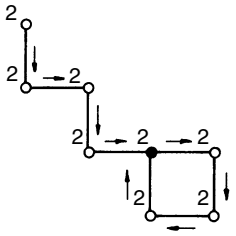


Fig. 3.8 A cluster and its graph showing how a loop is detected

To remove the problem of loops, we again make use of the trace. Each site which has been visited by the algorithm but was not burned (i.e., set to zero) receives the value 2. Hence a loop will be detected by testing whether the next site carries the value two. This site is then left alone and one can burn the previous sites.

The definition of the order parameter for the phase transition based on clusters is as follows: Take the largest cluster of a given configuration and divide it by the total number of occupied sites. Of course we have to average over many configurations to get a precise estimate of the order parameter P_∞.

Close to the transition point p_c we expect (Sect. 2.3)

$$P_\infty \propto |p - p_c|^\beta$$

with the exponent β. At this point we are not interested in how to compute β, but we want to determine p_c.

We have tacitly assumed so far that the boundaries are free. The lattice size is L^2 and once i or j is larger than L the boundary is crossed. The actual size of the lattice was assumed to be $(L+2)^2$ with the boundary set equal to zero. Here for the first time we encounter the issue of *boundary conditions*. There are two possible choices for the boundary. We can choose the boundary to be:

- Free or
- Periodic

For some problems even more complicated boundary conditions are useful. We will discuss these in the exercises to the section on importance sampling. Here we shall only be concerned with the free and periodic boundary conditions.

For the periodic boundary conditions we make the identification: $L + 1 = 1$ and $0 = L$. The result for the two-dimensional lattice is a torus. Do the results depend on the choice of boundary conditions?

▹ **Exercise 3.21**. Program the algorithm for the identification of clusters of the site percolation problem on the square lattice outlined above. Determine the order parameter P_∞ for various lattice sizes and probabilities p. Do this for the free and the periodic boundary conditions.

STOP

Here we encounter the finite-size problem again. We expect to find a power law behavior for the order parameter above the threshold and an identically zero order parameter below the threshold p_c. Instead, we find smooth S-shaped curves for all system sizes (Fig. 2.9).

In a finite lattice we cannot accommodate a diverging correlation length! The correlation length ξ diverges, when the transition point is approached, as

$$\xi \propto |p - p_c|^{-\nu}. \tag{3.3}$$

The correlation length can be at most $\xi \approx L$. To handle this situation, one applies a finite-size scaling analysis to the data for the different system sizes (Sect. 2.3.2). We shall defer such an analysis to the discussion of finite-size effects in the data for the Ising model. The only point we want to make now is that the simulation results depend on the choice of the boundary condition. The finite-size effects are different for free and for periodic boundary conditions!

For practical purposes the algorithm given above is of no use. For the large lattices one wants to study, the recursion goes so deep that the storage capacity will very soon be exhausted. In practice one uses a different algorithm which we discuss in the Appendix, where a C listing is also given.

Exercise 3.22 *Lattice animals*. Consider clusters on a regular lattice which have no loops. Such clusters are called lattice animals. Can you think of a way to use simple sampling for all lattice animals with N sites?

Exercise 3.23 *Anomalous diffusion*. Consider an ant parachuting onto a percolation configuration [3.8]. As soon as it has landed it starts a random walk on the clusters. The ant can only walk on the occupied sites. Compute the mean square displacement as a function of time and p:

$$\langle R^2 \rangle_p \propto t^x$$

and obtain the exponent x.

⋆ **Exercise 3.24**. Can you think of an algorithm for a parallel architecture which identifies clusters?

Variations on a theme

Exercise 3.25 *Bond percolation*. In the bond percolation problem one considers once again a lattice. For the site percolation problem all bonds between the sites were present. Only the sites mattered. Now, all the sites are occupied and we occupy a bond with a probability p and leave it empty with probability $1 - p$. Clusters are defined in a similar way as for the site problem. Determine the order parameter for the two-dimensional bond percolation problem. Do this problem for the order parameter defined by the spanning cluster P_s and for the order parameter determined by the mass of the largest cluster P_∞.

Exercise 3.26 *Site–bond percolation*. We can also study bond and site percolation together. Recall that for the site percolation problem all bonds existed and for the bond percolation problem all sites existed. In the site–bond percolation problem sites are present with a probability p_s. Bonds can exist only between occupied sites. If two neighboring sites are occupied the bond is present with a probability p_b. Study the percolation threshold p_{bc} as a function of p_s and p_b, in two dimensions.

Exercise 3.27 *Kinetic percolation*. For some problems we are interested not so much in all of the clusters in a given percolation configuration as in just the largest. For example, right at the percolation threshold p_c the largest cluster is a fractal and we would like to know more about its properties. We can generate a large percolation cluster by a sort of kinetic growth process from which we can also learn the growth law, i.e., how the radius of gyration grows with time. The algorithm is as follows. Start with a single site and make a list of the nearest neighbors. Occupy each of the sites in this list with probability p. Make a new nearest-neighbor list (be careful, the ones you visited before but decided not to occupy are marked as unoccupied). Program the algorithm and calculate the radius of gyration as a function of p and time.

Exercise 3.28 *Continuum percolation*. There is also a continuum version of the percolation problem. Imagine a square. You have disks at your disposal which you throw randomly onto the square. For this you generate an x coordinate and a y coordinate using a random number generator. Place the center of the disk at this coordinate with probability p. If you can cross from one side of the square to the opposite site via overlapping disks then a spanning cluster is found. Print out your trials and decide visually if there is a spanning cluster or not.

3.3 Biased Sampling

The treatment of the self-avoiding walk problem emphasized the limits of the simple sampling method. Though the method is fairly straightforward to apply, for many problems the usefulness of the concept is restricted. Even for a small number of steps it is hard to compile enough statistics. The difficulty increases exponentially with increasing number of steps. The concept of the biased sample helps (at least partially) to overcome this difficulty. To demonstrate how biased sampling can increase the efficiency, we stay with the self-avoiding walk problem. This also allows the reader to make an immediate comparison of the performance of simple sampling with that of the biased sampling method.

Algorithm 3.10 Self-avoiding random walk

```
do sample := 1 to n-of-samples
  begin
    step := 0;
    repeat
        generate-one-step;
        step := step + 1;
    until (step-invalid or step = N)
  end
  accumulate-results
end
```

3.3.1 Self-avoiding Random Walk

To begin to develop an algorithm for the biased sampling of the self-avoiding walk problem we look again at the basic algorithm we developed in the simple sampling treatment. This algorithm is Algorithm 3.10.

▹ **Exercise 3.29**. If you have missed out any of the exercises dealing with the simple sampling of self-avoiding random walks you should stop and do them now.

STOP

The real shortcoming of simple sampling was that a walk quickly terminates, i.e., the attrition problem. If we happen to choose a site which has been visited before, the walk terminates even though other sites are unoccupied. Suppose we always produce a list of the available nearest neighbors. The length of this list is always smaller than the coordination number of the lattice minus 1, i.e., $2d - 1$, on a d-dimensional hypercubic lattice. Let l be the length of the list of available neighbors. In the biased sampling we choose only from the list of available neighbors, and each entry has the same probability of being picked. The probability of each entry is hence $1/l$.

The idea of picking only the available sites does not guarantee that the walk continues until N steps have been made. The problem that the walk runs into a dead end before N steps have been made still remains, but the probability that the walk continues up to N steps is increased.

▹ **Exercise 3.30**. Program Algorithm 3.11 for the biased sampling of self-avoiding random walks. Print out the number of steps where the walks terminated.

STOP

For this algorithm, as before for simple sampling of the nonreversal and self-avoiding random walks, labels are a useful tool to keep track of the path and to generate a list of valid neighbors.

Algorithm 3.11 Self-avoiding random walk

```
integer lattice(−N : N, −N : N);
do sample := 1 to n-of-samples
  begin
    step := 0;
    repeat
      produce-valid-neighbor-list(x, y);
      if list = empty then
        terminate := true
      else
        begin
          choose-new-site-from-list(x_new, y_new);
          x := x_new; y := y_new;
          step := step + 1;
        end
    until (terminate or step = N);
  accumulate-results
end
```

Those who did Exercise 3.30 will have found that the lengths of the walks we could sample before a walk terminated has increased. The price we pay for this improvement is that we have introduced a *bias*. We are favoring certain configurations of a walk. For simple sampling, every walk carried the same weight for average quantities like the end-to-end distance. For biased sampling every walk of length N steps carries a weight

$$W = \prod_{i=1}^{N} \frac{l_i}{(2d-1)}, \tag{3.4}$$

and when calculating averages this has to be taken into account.

▹ **Exercise 3.31**. Calculate the average end-to-end distance with the biased sampling of self-avoiding random walks. Compare your results with those from simple sampling.

STOP

Variations on a theme

! **Exercise 3.32**. *k-tuple SAW*. The k-tuple self-avoiding random walk is very similar to the regular self-avoiding random walk. Whereas in the regular SAW a site can be visited at most once, the k-tuple walk is allowed to visit a site at most k-times. Write an algorithm for this problem using the biased sampling concept.

Exercise 3.33 *GSAW*. The growing self-avoiding random walk (GSAW) is defined almost the same as the regular self-avoiding walk. The only difference is that the transition probability to a site is the inverse of the number of free sites. Write an algorithm for this problem using the biased sampling concept.

3.4 Importance Sampling

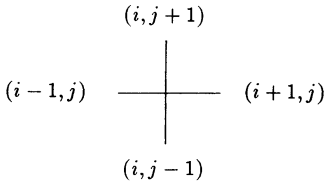
The concept of simple or biased sampling fails to be practical for problems where contributions to quantities come almost entirely from certain regions in phase space. A homogeneous sampling of the phase space would require a tremendous effort. The knowledge of the regions where the important contributions come from can be used to sample mainly in these regions. A typical example for the application of this idea is the Ising model.

3.4.1 Ising Model

As with the previous examples we will concentrate on studying the problem in a two-dimensional space. Consider the Ising model Hamiltonian

$$\mathcal{H} = -J \sum_{\langle i,j \rangle} s_i s_j, \tag{3.5}$$

where J (the exchange interaction between nearest-neighbor sites) can be either positive, i.e., the ferromagnetic case, or negative, the antiferromagnetic case. The variable s denotes a spin of either $+1$ or -1. We shall restrict ourselves to a simple square lattice for which the symbol $\langle i, j \rangle$ means the four neighbors nearest to the site (i, j):



Let $\boldsymbol{x}$ denote a configuration of spins.

Recall, from the discussion on the importance sampling concept in Chap. 2 and at the beginning of this chapter, that we want to generate a *Markov chain*

$$\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$$

such that the configuration $\boldsymbol{x}_{i+1}$ depends only on the immediately preceding configuration $\boldsymbol{x}_i$. The probability of getting to $\boldsymbol{x}_{i+1}$ from $\boldsymbol{x}_i$ is given by a *transition probability*

$$W(\boldsymbol{x}_{i+1}|\boldsymbol{x}_i).$$

Why do we need a Markov chain? The idea of importance sampling is to sample mainly the largest contributions to a quantity, for example, the magnetization. We want to sample configurations of spins where the Boltzmann factor $\exp[-\mathcal{H}(\boldsymbol{x})/k_\mathrm{B}T]$ has a peak. A priori we do not know where such regions are! Suppose we could invent a process by which states are generated where an important contribution to the magnetization or any other observable is made. This process is the Markov process, with an appropriate transition probability leading from one configuration to another.

The transition probability is chosen such that ultimately the distribution of the states $\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ is the *Boltzmann distribution*

$$P(\boldsymbol{x}) \propto \exp\left(-\frac{\mathcal{H}(\boldsymbol{x})}{k_\mathrm{B}T}\right).$$

How must we choose the transition probability to guarantee that the states will be distributed with a Boltzmann law? We have already given a detailed discussion of this point in Sects. 2.1 and 2.2. Intuitively we can argue as follows. Suppose we restrict the treatment to attempts to turn over single spins s_i. Suppose on turning a spin from s_i to $-s_i$ we lose energy. Because we always want to be at or near the ground state of the model we should accept such a move with probability one. Hence, in the case where the change in energy $\Delta\mathcal{H}$ is negative, $\Delta\mathcal{H} = \mathcal{H}(\boldsymbol{x}_\mathrm{new}) - \mathcal{H}(\boldsymbol{x}_\mathrm{old})$, we have $W(\boldsymbol{x}_\mathrm{new}|\boldsymbol{x}_\mathrm{old}) = 1$. However, this way we will certainly get stuck in a local minimum of the energy. To avoid this we accept moves which raise the energy. But we allow moves which raise the energy much only very rarely; they should have a low probability. If, on the other hand, the change in the energy is small, i.e., the energies of the old and new configurations are close, we accept the move with a fairly high probability. This way we can climb out of a local energy minimum.

How should we choose the transition probability in the case where the change in energy $\Delta\mathcal{H}$ is positive? Intuitively it seems reasonable, and it can be shown exactly (Sects. 2.1 and 2.2), that we can choose

$$W(\boldsymbol{x}_{i+1}|\boldsymbol{x}_i) = \min\left[1, \exp\left(-\frac{\Delta\mathcal{H}}{k_\mathrm{B}T}\right)\right], \tag{3.6}$$

which is called the *Metropolis function*. This is, however, not the only possible choice for a transition probability. There are other choices, for example the *Glauber function* $(\frac{1}{2}(1 - \tanh(\Delta\mathcal{H}/2k_\mathrm{B}T))$.

▹ **Exercise 3.34**. What are the possible changes in energy $\Delta\mathcal{H}$ for the two-dimensional nearest-neighbor Ising model on a simple lattice?

STOP

Above we constructed a *single-spin-flip* transition probability. Once all spins have been given a chance to reverse their directions one *sweep* has been made. One sweep is also called one *Monte Carlo step per spin*, abbreviated MCS.

There are several possible ways to implement a sweep through the lattice. Taking the word sweep literally, we mean going regularly through the lattice. Another possibility is to draw the site to be visited at random. Once as many attempts as there are sites in the lattice have been made, one Monte Carlo step has been made. In one Monte Carlo step some sites may have been considered several times, whereas others were completely ignored. On the average, barring any problems with the random numbers (Sect. 3.2.1), each site is given a chance for a reversal.

Now that we know how to set up the Markov process to sample the important contributions to an average, how do we compute the magnetization, for example? Due to the construction (Sect. 2.1) the computation of the magnetization appears as a simple average. The magnetization of a configuration is given by

$$M_i = \sum_{j=1}^{L^2} s_j \tag{3.7}$$

and the magnetization per spin is then

$$m_i = M_i L^2. \tag{3.8}$$

The *average* magnetization per spin is the arithmetic average over many configurations

$$\langle m \rangle = \frac{1}{N} \sum_{i}^{N} m_i. \tag{3.9}$$

We have to start the importance sampling with some initial configuration $\boldsymbol{s}_0$ and from then on the new states are generated. Which initial spin configuration should we choose? The answer depends on the type of problem we are interested in. Let us assume that we want to study the equilibrium properties of the model, such as the magnetization and susceptibility, at certain temperatures. Suppose, further, that we do not start from a completely ordered spin configuration but from a random spin configuration. What will happen? The random spin configuration corresponds to the infinitely high temperature state $T = \infty$ with zero spontaneous magnetization. The temperature we are interested in is below the critical temperature so that there is a nonzero spontaneous magnetization. We now let the system evolve in time using, for example, the Metropolis function. Essentially, we have performed a quench from a disordered region into the ordered region in the phase diagram. It will take an enormously long time until an equilibrium state is established. Before equilibrium is established the excess energy has to be removed from the configuration.

Algorithm 3.12 Ising model

```
(* initialize the lattice *)
  do i := 1 to L
    do j := 1 to L
      lattice(i, j) := −1
  (* Monte Carlo part *)
  do mcs := 1 to mcsmax
    begin
      generate-one-sweep;
      if mcs ≥ n0 then
        do-analysis
    end
```

To avoid this problem we start out with an ordered configuration, and then the initial transient in the Markov chain, which corresponds to nonequilibrium states, is much shorter. Away from the transition point the relaxation is exponentially fast. What happens close to the transition point has been discussed in Sect. 2.3. For convenience we shall start with a configuration where all spins are down, i.e., with a magnetization of -1. Alternatively we could always start from the last configuration generated for some temperature if we go up in temperature.

We are almost ready to write part of the algorithm. What we have to consider is the sample size. For the moment let us leave this question on one side and call the number of samples we want to take *mcsmax*.

In the Algorithm 3.12 n_0 is the number of configurations of the transient which we must discard because they do not correspond to equilibrium configurations for the desired temperature.

Let us go back again to the transition probabilities. It is clear that the configuration $\boldsymbol{x}_{i+1}$ cannot differ too much from the configuration $\boldsymbol{x}_i$. We have allowed only single spin flips and not flips of larger patches of spins, which would lead to a larger stride in phase space. The sequence of the states can be viewed as a trajectory in phase space (see discussion in Sect. 2.2.2 on the dynamic interpretation of the Monte Carlo algorithm). Viewed from this angle we can see that successive configurations must be correlated. To obtain an unbiased statistical sample of states we must discard n_δ states between two sample points. How many we must discard depends not only on the temperature but also on the conservation laws. See Algorithm 3.13.

Exercise 3.35 Can you design another solution to select every n_δth configuration?

After these preliminaries we come to the central part of the importance sampling algorithm of the Ising model, i.e., to *generate-one-sweep* in Algorithm 3.13. The transition probability was designed such that we always attempt to flip only a single spin and not larger patches of spins. One sweep through the lattice is to attempt to flip all spins. This is one *Monte Carlo step*. After one Monte Carlo step we have obtained a new configuration from the old. So the first step is Algorithm 3.14, where L is the linear lattice size.

Algorithm 3.13 Ising model

```
(* initialize the lattice *)
  do i := 1 to L
    do j := 1 to L
      lattice(i, j) := −1
    (* Monte Carlo part *)
  count := 0;
  do mcs := 1 to mcsmax
    begin
      generate-one-sweep;
      if mcs ≥ n0 then
        begin
          count := count + 1;
          if count = nδ then
            begin
              count := 0;
              do-analysis;
            end
        end
    end
```

Algorithm 3.14 *generate-one-sweep*

```
do i := 1 to L
  do j := 1 to L
    begin
      compute-the-energy-difference;
      decide-to-flip-or-not
    end
```

Algorithm 3.15 *generate-one-sweep*

```
do i := 1 to L
  do j := 1 to L
    begin
      compute-the-energy-difference;
      if random(iseed) < W(ΔH) then
        flip-the-spin;
    end
```

Recall from the above discussion that we want to accept a spin flip with a probability W, which we take equal to the Metropolis function. Suppose we have computed the energy difference $\Delta\mathcal{H}$. We can evaluate the probability W of accepting the move because it depends only on this difference $W(\Delta\mathcal{H})$.

Choose a random number between 0 and 1. If the random number is less than the computed transition probability we should flip the spin. Otherwise we should leave the spin orientation as it was (Algorithm 3.15).

Algorithm 3.16 *look-up-table*

```
real W(−4 : 4);
do j := −4 to 4 step 2
  begin
    W(j) := 1;
    if j > 0 then
      W(j) := exp(−2 ∗ Jkt ∗ j)
  end
```

We realize that there is no need to evaluate the function W every time we want to flip a spin. There are only 5 possible values for the energy difference (Exercise 3.34). The function can be precomputed and stored in an array for easy look-up. The energy difference is given by

$$\Delta\mathcal{H} = 2s_i \frac{J}{k_B T} \sum_{\mathrm{nn}(i)} s_j, \tag{3.10}$$

where the symbol $\mathrm{nn}(i)$ denotes the 4 nearest neighbors of the central spin i.

There is only one problem left. We have to resolve the last heading where we compute the energy difference. Here there is a branching point. We are faced with the same decision as for the percolation problem. We must choose between different boundary conditions. The bonds of every spin at the edges of the lattice cannot be left dangling in the air. The spins must interact with nearest neighbors. For now we elect to use the *periodic boundary conditions*.

One way to implement the periodic boundary conditions is to use the *modulo* function. At the edges of the lattice, i.e., where for the central spin (i, j) either i is 1 or L or j is 1 or L we compute the nearest neighbors by

$$ip1 = i + 1 \pmod{L} \tag{3.11}$$
$$im1 = i - 1 \pmod{L} \tag{3.12}$$
$$jp1 = j + 1 \pmod{L} \tag{3.13}$$
$$jm1 = j - 1 \pmod{L}. \tag{3.14}$$

This approach is very time consuming. Each evaluation of the *modulo* function involves several arithmetic operations. Time is crucial. The innermost loop has to be constructed so as to minimize the number of arithmetic operations. The solution here is again to use a table. Actually we will need two tables. One for the "plus" operation and one for the "minus" operation (Algorithm 3.17).

Now the algorithm to perform an importance sampling for the two-dimensional Ising model is complete. All major parts have been worked out. For completeness we give the full algorithm as Algorithm 3.18.

Algorithm 3.17 *look-up-table-for-mod*

```
integer ip(1 : L), im(1 : L);
do i := 1 to L
  begin
    ip(i) := i + 1;
    im(i) := i − 1;
  end
ip(L) := 1;
tm(1) := L;
```

▹ **Exercise 3.36**. Program Algorithm 3.18. Choose some values for the parameter *Jkt* and print out the evolution of the magnetization as a function of the Monte Carlo steps.

STOP

Our premier goal is to determine the phase transition point, i.e., $J/k_B T_c$, for the two-dimensional Ising model on the simple lattice. Of course, the analytic result for the two-dimensional Ising model is known exactly [3.9]. This gives us the opportunity to see how well the data analysis works.

In the discussion in Sect. 2.3 we learned that the phase transition point is smeared out for systems of finite size. These finite-size effects depend on the boundary conditions. The results for the order parameter and the susceptibilities for simple sampling of the percolation problem turned out to be different. The same is true for thermally driven phase transitions.

How can we determine the phase transition point when the data depend on the system size as well as on the boundary conditions? The key to the solution of this question is the *scaling* feature of the system near the transition point. Close to the transition point we expect to find the scaling relations

$$m(T, L) = L^{-\beta/\nu} \hat{f}\left[(T - T_c)L^{1/\nu}\right], \tag{3.15}$$

$$\chi(T, L) = L^{\gamma/\nu} \hat{g}\left[(T - T_c)L^{1/\nu}\right]. \tag{3.16}$$

The scaling functions $\hat{f}$ and $\hat{g}$ depend on the details, i.e., on the boundary conditions.

▹ **Exercise 3.37**. Determine the order parameter and the susceptibility from the order parameter fluctuations as a function of the temperature T and the linear system size L. Do the simulations for free and for periodic boundary conditions. Evaluate also the fourth moment of the magnetization.

STOP

Having obtained the raw data from Exercise 3.37 we can now analyze these using the ideas of *finite-size scaling*. An example of such raw data is shown in Fig. 3.9.

Algorithm 3.18 Ising model

```
integer lattice(1 : L, 1 : L);
integer ip(1 : L), im(1 : L);
real W(−4 : 4);
  (* look-up-table-for-mod *)
do i := 1 to L
  begin
    ip(i) := i + 1;
    im(i) := i − 1;
  end
ip(L) := 1;
im(1) := L;
  (* look-up-table *)
do j := −4 to 4 step 2
  begin
    W(j) := 1;
    if j > 0 then
      W(j) := exp(−2 ∗ Jkt ∗ j)
  end
  (* initialize the lattice *)
do i := 1 to L
  do j := 1 to L
    lattice(i, j) := −1;
  (* Monte Carlo Part *)
count := 0;
do mcs := 1 to mscmax
  begin
  do i := 1 to L
      do j := 1 to L
        begin
          ici := lattice(i, j);
          ien := lattice(ip(i), j) + lattice(im(i), j)+
            lattice(i, ip(j)) + lattice(i, im(j));
          ien := ici ∗ ien
          if ranf < W(ien) then
            lattice(i, j) := −ici
        end
      if mcs ≥ n0 then
        begin
          count := count + 1;
          if count = nδ then
            begin
              count := 0;
              do-analysis;
            end
        end
  end
```

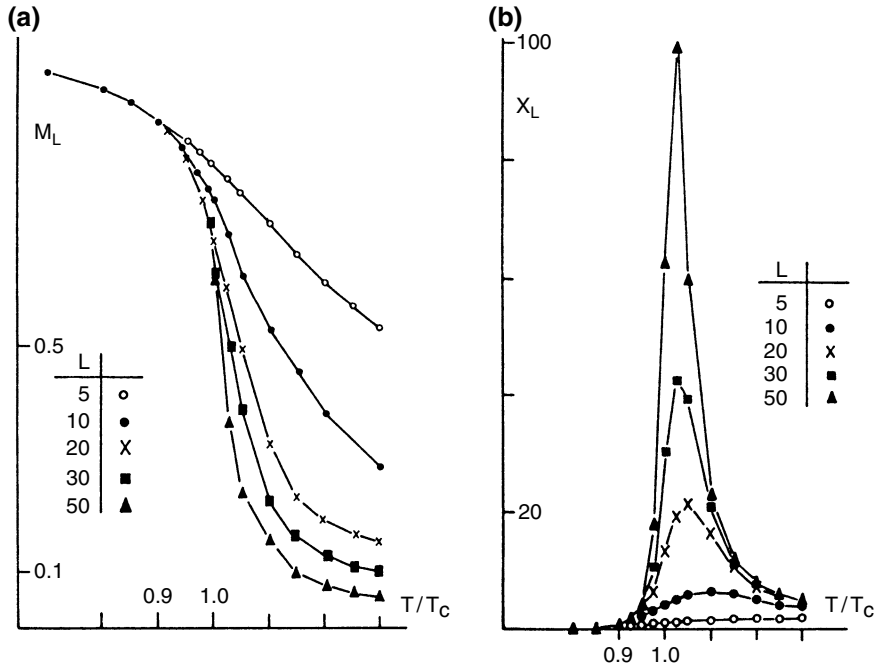


Fig. 3.9 Dependence of the magnetization and susceptibility on the linear lattice size L. The data shown are for the two-dimensional Ising model

Before actually doing the finite-size scaling analysis we should reflect on the accuracy of the raw data. Were there enough initial configurations discarded? Recall that close to the transition point the system relaxes into equilibrium with a characteristic time τ

$$\tau \propto L^z,$$

where z is of the order of 2. Also, the configurations are dynamically correlated. Recall that the dynamic correlation between the configurations reduces the sample size (Sect. 2.3.8).

Before carrying on, can you answer the following question in the affirmative?

Did you accumulate the *absolute* value of the magnetization?

If the answer to the question is *no* you have to repeat all your simulations! But why?

The reason for doing so has been discussed in Sect. 2.3. Due to the finiteness of the system we find that it is ergodic below the transition point. In an infinite system, i.e., in the thermodynamic limit, there is no chance for the system to go from one branch of the spontaneous magnetization to the other. For systems of finite size there is a characteristic time, which depends on the system size, in which the system can manage to cross over to the other branch. Those who carefully did the exercises, in

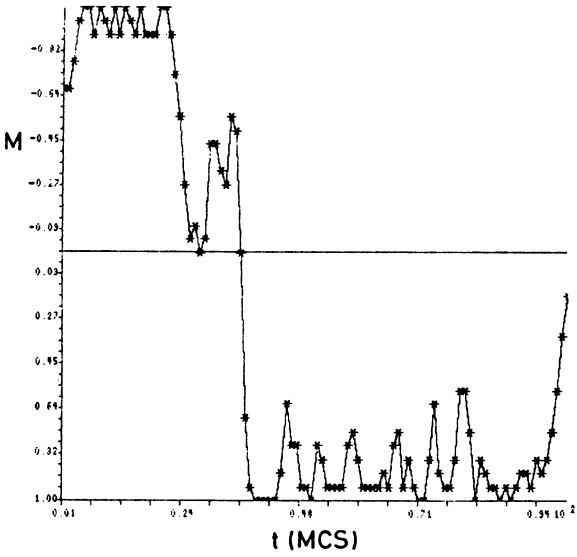


Fig. 3.10 Evolution of the magnetization as a function of the Monte Carlo steps (MCS). This figure shows that for small systems the magnetization changes branches quite often

particular Exercise 3.36, should have noticed this phenomenon. Starting out from the negative side, all spins were down. Then the system relaxed into equilibrium and fluctuated around the equilibrium. Sometimes the magnetization changed to positive values for temperatures below the transition point. An example where this happened is shown in Fig. 3.10.

Simply averaging the magnetization would yield an incorrect value. It depends on the number of times the system reversed its magnetization. To avoid the problem the two branches are folded on top of each other by taking the absolute value of the magnetization.

After this long discussion of the do's and don'ts we can go ahead and determine the transition point. For this we use the cumulant

$$U_L = 1 - \frac{\langle m^4 \rangle_L}{3\langle m^2 \rangle_L^2}. \tag{3.17}$$

There is little extra cost in computing this quantity in the simulations. Here we have to issue a warning. Quite often the data are such that a variable with single precision is not sufficient for the accuracy needed. Valuable digits are lost using a single precision variable which render the results absolutely useless. This point cannot be emphasized enough.

To determine the transition point we choose pairs of linear system sizes (L, L'). The critical point is the fixed point where we have

$$U_L = U_{L'}. \tag{3.18}$$

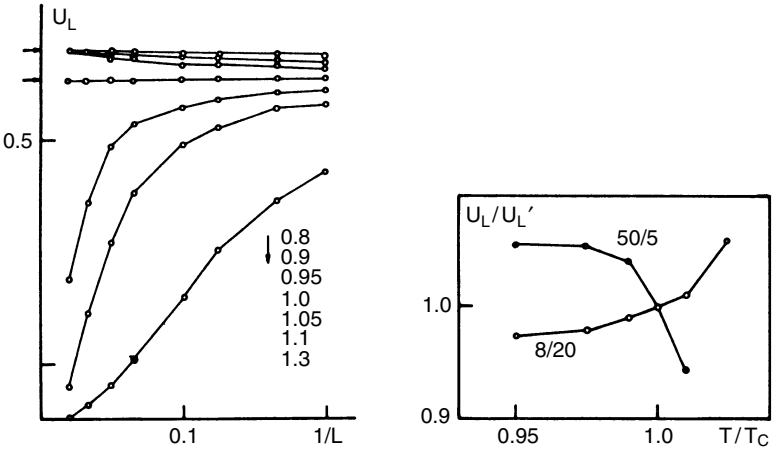


Fig. 3.11 Analysis of the cumulants for the two-dimensional Ising model

At the transition point the cumulants for two different system sizes must intersect. To put it differently, if we plot the ratio $U_L/U_{L'}$ versus the temperature, then for all pairs we should find a unique crossing at one particular temperature. This is the transition point.

▹ **Exercise 3.38**. Carry out the cumulant analysis to determine the critical point of the two-dimensional Ising model. Compare your answer to the exact value $J/k_B T_c = \frac{1}{2}\ln(1+\sqrt{2})$.

STOP

Figure 3.11 shows an analysis of the cumulants for the two-dimensional Ising model.

In practice the analysis of the data on the cumulants is not as straightforward as it seems. To obtain a unique intersection point for all pairs (L, L') one has to work fairly close to the expected critical point. This is usually an iterative process. A first scan of the temperature range narrows down the interesting region by inspection of the magnetization and the susceptibility. In the suspected critical region points are selected for further simulation runs. For these points one needs very good statistics. With the data a first analysis of the cumulants is made. It often happens that a unique intersection cannot be found at this stage. More points with even better statistics are needed.

Now that we have determined the phase transition point for the model we can go on to the *finite-size scaling analysis*. All the data on the magnetization for the different system sizes can be made to collapse onto a single scaling function. The scaling function depends on the boundary conditions.

The scaling depends also crucially on the exponents and on the transition point. If the transition point was not determined precisely enough we will find a bad scaling behavior, if we find scaling at all. If the exponents are not known precisely enough then this will also show up in the scaling behavior.

Fortunately, for the two-dimensional Ising model we know the exponents exactly. We also know the transition point exactly. We use this precise knowledge to see how well the data obtained in Exercise 3.37 scale. If the scaling turns out to be poor, then this reflects problems with the accuracy of the data.

For the scaling we plot on the x axis the variable

$$x \equiv (T - T_c)\, L^{1/\nu}$$

and on the y axis we plot the variable

$$y \equiv m(T, L) L^{\beta/\nu}.$$

The data for the different system sizes should, when the correct values for the parameters T_c, β and ν are chosen, collapse onto a single curve. The converse statement is also true. If upon varying the parameters T_c, β and ν all the data collapse onto a single curve then we have found the critical point and the critical indices.

▹ **Exercise 3.39**. Carry out the finite-size scaling analysis for the magnetization.

STOP

To conclude this section we take up again the discussion on the accuracy of the data. We learned in Sect. 2.3.8 that certain quantities are not *self-averaging*. Upon going to larger system sizes the error in these quantities does not decrease. The error is independent of the system size.

One such quantity exhibiting a lack of self-averaging is the susceptibility. We gain nothing in accuracy by going to larger system sizes, we only have to work harder.

▹ **Exercise 3.40**. Confirm numerically that the susceptibility exhibits a lack of self-averaging.

STOP

! **Exercise 3.41**. Carry out the finite-size scaling analysis for the susceptibility.

! **Exercise 3.42**. *Glauber function*. At the beginning of this section we introduced the Glauber function as another possible choice of a transition probability employing a single spin flip. Incorporate the Glauber function into your existing program for the two-dimensional Ising mode. Redo some of the exercises in this section.

! **Exercise 3.43**. The Hamiltonian of the Ising model in a magnetic field H is given by

$$\mathcal{H} = -J \sum_{\langle i,j \rangle} s_i s_j - H \sum_i s_i .$$

What are the possible changes in the energy $\Delta\mathcal{H}$ for a single spin flip? Redo Exercise 3.37.

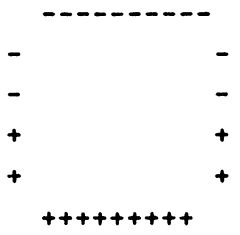
Exercise 3.44 Invent a mapping which maps the two-dimensional simple lattice to one dimension, such that all neighbors are evenly spaced. Use periodic boundary conditions.

Exercise 3.45 Implement helical (= skew) boundary conditions.

Exercise 3.46 *Checker-board algorithm*. To efficiently run a program on a computer with a pipeline architecture the algorithm must be in a vectorizable form. For the vectorization of an algorithm for the two-dimensional Ising model it is useful to split the lattice into two sublattices. There are several ways of doing so. One way is to color the sites black and white like a checker board. The energy change on reversing, say, a black site depends only on the white sites surrounding this site. All black sites can be updated simultaneously! Write an algorithm for the two-dimensional Ising model which employs the checker board idea.

Exercise 3.47 What problem corresponds to the limit $T = \infty$ in the Ising model?

Exercise 3.48 *Interfaces*. So far the boundary conditions we have employed have been homogeneous. The boundary was either periodic or all sites of the boundary were constrained to one fixed value. Consider a boundary like



Such a boundary produces an interface in the system. Use the program you developed in this section and set up the boundary as shown above. Print out the configurations you obtain.

Exercise 3.49 *Kawasaki dynamics*. The way we have treated the importance sampling of the Ising model has been such that there was no constraint, beside the chosen temperature, placed on the system. Sometimes we want to introduce a conservation law. For example, we want the concentration of the $+1$ and the -1 spins to be fixed. To realize a fixed concentration we introduce instead of single spin flips an exchange between spins. Pick a pair of nearest-neighbor spins with different orientation and exchange them. Compute the change in the energy and proceed as for the single spin flip, i.e., if the move is rejected the pair is not exchanged. What are the possible values for the change $\Delta\mathcal{H}$ in the energy?

Variations on a theme

Exercise 3.50 *Potts model*. The Hamiltonian of the Potts model is given by

$$\mathcal{H}_{\text{Potts}} = -\sum_{\langle i,j\rangle} J_{ij}\delta_{q_i q_j}.$$

Here J_{ij} are exchange couplings which can be all identical. The variable q can take on the values $\{1, 2, \ldots, q\}$. Determine the critical point for the two-state Potts model.

Exercise 3.51 *XY model*. The XY model is an example of a model with continuous degrees of freedom. There are no spins which can take on only discrete values. One way to write the Hamiltonian for the XY model is

$$\mathcal{H}_{XY} = -\sum_{\langle i,j\rangle} J_{ij}\cos(\phi_i - \phi_j) - H\sum_i \cos\phi_i.$$

Here J_{ij} are exchange couplings which can be all identical; H is a magnetic field. Write an algorithm for importance sampling of the XY model.

Exercise 3.52 *Lennard–Jones*. All sections have dealt entirely with examples where a lattice was at least partly involved. This, of course, is not necessary for the application of the importance sampling idea. Consider N point particles in a box of volume V. The volume is chosen such that the simulation can be carried out for a particular density ϱ. The particles interact with each other via a Lennard–Jones pair potential

$$\Phi(r_{ij}) = 4\varepsilon\left[(\sigma/r_{ij})^{12} - (\sigma/r_{ij})^{6}\right].$$

Here, r_{ij} is the distance between particle i and j; ε specifies the unit of energy and σ the unit of length. It is most convenient to work with scaled variables. An importance sampling for the Lennard–Jones system proceeds as follows. Choose a particle. Generate a random number r and displace the particle a distance δ derived from the random number. Calculate the change in the energy. The acceptance of the displacement is done completely analogously to the acceptance discussed in the text. Write an algorithm and program it. Compute the average internal energy, specific heat and pressure. How do you incorporate the boundary conditions?

3.4.2 Self-avoiding Random Walk

To close the guide what could be more appropriate than an exercise:

▹ **Exercise 3.53**. *Kink-jump*. Consider the self-avoiding random walk on a simple square lattice. We draw on the analogy with polymers and view the self-avoiding walk as a chain of N beads and rods (or bonds and sites if you wish). To generate new configurations pick a bead (or site) at random. If the bead is an end bead it could, in principle, be moved in three directions. If the bead is not an end bead a "kink" type motion could be made:

1. Plot a few examples of conformations of a chain and list the possible motions of the beads.
2. Invent an algorithm for the kink-jump method. What kind of sampling are you doing?
3. Do you see ergodicity problems?
4. So far we have not considered interactions between the beads. Suppose there is an interaction ε between beads similar to the exchange interaction in the Ising model. Invent an algorithm for a simulation of this model. What kind of sampling are you doing?

Part II
Survey of More Advanced Simulation Methods

Chapter 4
Cluster Algorithms and Reweighting Methods

4.1 Introduction

Roughly at the time (1987) when the manuscript for the first three chapters of the present book was completed, several breakthroughs occurred. They had a profound influence on the scope of Monte Carlo simulations in statistical physics, particularly for the study of phase transitions in lattice models.

The first of these remarkable developments is the invention of the "cluster algorithms" [4.1–4.54] that reduce (or eliminate completely) the problem of "critical slowing down" (divergence of relaxation times near the critical point, see e.g., (2.85), (2.86) [2.59]) from which the single spin-flip algorithms suffer, since the increase of relaxation time means a dramatic increase of statistical errors; see, e.g., (2.87). The original version [4.1], proposed for the Ising model, is based on the mapping [4.55] between Potts models [2.106] and percolation [2.60], but meanwhile extensions exist to a wide variety of models, including also isotropic magnets [4.7, 4.9] and various quantum spin systems [4.54]. In this chapter, only the original version (for Ising and Potts problems) will be considered, however. Here we briefly outline the main physical idea: critical slowing down in the kinetic single-spin-flip Ising model [2.12] can be attributed to the fact that the long-range critical spin correlations correspond to the occurrence of large clusters of correlated spins. It takes a very long time until such a cluster disintegrates and finally disappears by many subsequent single spin flips. However, if the basic move is not a single spin flip but the overturning of a whole cluster, the cluster pattern changes rapidly and one thus can move much faster through phase space, even close to the critical point. Of course, the "art" is to construct the clusters that are considered for flipping such that at the same time one has a high acceptance rate for such a move, even if the cluster is very large, and that one destroys physical correlations by the successive action of the algorithm. There is no general recipe for how one can construct such an algorithm for a particular model – in some cases, where critical slowing down is very dramatic, like in spin glasses [2.41], intensive searches for efficient cluster algorithms have been made but have so far failed.

K. Binder and D. W. Heermann, *Monte Carlo Simulation in Statistical Physics*,
Graduate Texts in Physics, https://doi.org/10.1007/978-3-030-10758-1_4

Another line of research, based on the idea that Monte Carlo updates should reduce the problem of critical slowing down if they operate on all length scales simultaneously, not only on the length scale of the lattice spacing as the single spin-flip-type algorithms do, is known as "multigrid Monte Carlo" [4.56–4.61]. However, often the asymptotic behavior does not change qualitatively: i.e., the dynamic exponent z in the relation between relaxation time τ and linear dimension L, $\tau \propto L^z$ (2.86), remains the same as in the single spin-flip-type algorithm. But the constant of proportionality in this relation may get markedly smaller (a decrease by a factor of 10 was often reported [4.60]). Nevertheless, we shall not deal with this approach here further, noting that it also lacks another advantage that the "cluster algorithms" have, namely that the latter yield in a very natural way "improved estimators" for various quantities: expressing quantities like susceptibilities [4.7, 4.96], pair correlation functions [4.7, 4.25], fourth-order cumulants [4.62], etc., in terms of "clusters" rather than using the original formulation in terms of spins, one exploits the fact that there are no correlations between different clusters, and hence the statistical noise is reduced.

The second important development that we shall discuss is the reanimation [4.63] of the old idea [4.64, 4.65] of "reweighting": from a simulation at a single state point (characterized in an Ising model by choice of temperature T and magnetic field H) one does not gain information on properties precisely at that point only, but also in the neighboring region. Extensions and variations of this concept are indeed very promising and powerful [4.46, 4.66–4.69]. The simplest approach, the "single histogram method" [4.63], starts from the observation that the distribution of the energy $P(E, T)$ at temperature T can be obtained from the distribution $P(E, T_o)$ at a neighboring temperature T_o by

$$P(E,T) = P(E,T_o)\exp[-(1/T - 1/T_o)E/k_B]/\sum_E P(E,T_o) \times \exp[-(1/T - 1/T_o)E/k_B]. \tag{4.1}$$

Of course, in this simple form the method is useful only either for rather small systems [where $P(E, T_o)$ is sufficiently broad since the width of the distribution scales with $L^{-d/2}$, assuming a d-dimensional cubic box of linear dimension L as the simulation volume] or at a critical point, as first recognized by *Ferrenberg* and *Swendsen* [4.63]: due to the critical fluctuations, the distribution is broadened over a width of order $L^{-1/\nu}$, where ν is the critical exponent of the correlation length. Since in a finite-size scaling analysis, see (2.79) and (2.80), the region of interest is of the order of $|1 - T/T_c| \propto L^{-1/\nu}$, the region of interest in finite-size scaling is of the same order as the region in temperature accessible for reweighting. Of course, there still is a problem in the wings of the distribution, where due to reweighting statistical errors may get greatly magnified, but this problem is eased by the combined use of several histograms at suitably chosen neighboring temperatures (or other control parameters), the so-called "multiple histogram extrapolation" [4.46, 4.66, 4.70].

Particularly interesting are also reweighting schemes built into the simulation procedure: again this is an old idea under the name of "umbrella sampling" [4.71–4.73], but gained full strength in a version of the technique which is called "multicanonical Monte Carlo" [4.67, 4.74] or "multimagnetic ensemble" [4.75–4.77] (reweighting with respect to magnetization rather than energy). These techniques have now become standard tools for the study of asymmetric first-order transitions, interfacial free energies, etc., when one combines them suitably with finite-size scaling ideas, as will be discussed below.

A related approach ("simulated tempering" [4.78], "expanded ensemble" [4.79]) considers the temperature just as an additional degree of freedom, specifying properly transition rules for jumping from one temperature to another one. In this context, we note also the "broad histogram method" and related methods [4.80] which focuses on sampling the density of states $g(E)$ directly.

In this chapter, we cannot attempt to give a full account of all these techniques (or further variations to this theme [4.81–4.84]). Thus, in Sect. 4.3 we shall describe a "case study" which nicely illustrates the power of the approach, but also hints to some of the problems that have to be taken care of. Similarly, Sect. 4.2 describes another "case study" where the Swendsen–Wang cluster algorithm [4.1] is used in conjunction with finite-size scaling techniques. Finally, Sect. 4.4 discusses some of the advances that recently have been made in the application of finite-size scaling methods themselves, in particular the concept of "field mixing" [4.85–4.94] that has enabled the study of critical phenomena in asymmetric systems such as liquid–gas transitions [4.85–4.89], unmixing of polymer solutions [4.90], etc. Also the problem of "crossover" [4.95] from one "universality class" [4.96] to another will be mentioned and possible treatments by simulation methods indicated [4.97–4.101].

4.2 Application of the Swendsen–Wang Cluster Algorithm to the Ising Model

If the reader compares the finite-size scaling description for the Ising model, e.g., (2.72) and (2.80), to the corresponding results for the percolation problem, (2.63) and (2.64), he will note a complete formal analogy (of course, the explicit form of the scaling functions and the critical exponents of both problems must differ, since these problems constitute different universality classes [4.96], but the general structure is the same). This analogy is no surprise at all, since the mapping proved by *Fortuin* and *Kasteleyn* [4.55] implies that bond percolation is equivalent to the limit $q \to 1$ of the q-state Potts model [2.106], cf. (2.106) for a definition of this model (note that the case $q = 2$ is nothing but the Ising model). As a consequence, the thermal order–disorder transition of the Ising model (and related spin models) can be described as a percolation of "physical clusters" [4.102, 4.103].

Of course, it has been known for a long time that any state of the Ising lattice can be described in terms of "geometrical clusters" of, say "down spins" in a surrounding background of "up spins" [4.104, 4.105]. However, throughout the paramagnetic phase we encounter a percolation transition of these "geometrical clusters" when we vary the magnetic field H from strongly positive to negative values. The clue to identify "physical clusters" in contradistinction to "geometrical clusters" is the concept of "active bonds" [4.102]. No such bonds can occur between spins of opposite sign, while for a pair of neighboring spins the probability p for a bond to be active is

$$p = 1 - \exp(-2J/k_{\mathrm{B}}T)\,, \tag{4.2}$$

and only spins connected by "active" bonds form a "physical cluster". Actually (4.2) can be extended to ferromagnetic interactions of arbitrary range [4.51], not restricted to nearest neighbors, and very efficient cluster algorithms have been developed for such long-range Ising models [4.51], but this is out of consideration here.

This interpretation, as stated above, can be rigorously proven by the percolation representation of the partition function of the Potts model [4.55], i.e.,

$$\begin{aligned} Z_{\mathrm{Potts}} &= \operatorname*{Tr}_{\{\sigma_i\}} \exp(-\mathcal{H}_{\mathrm{Potts}}/k_{\mathrm{B}}T) \\ &= \sum p^{N_{\mathrm{b}}}(1-q)^{N_{\mathrm{m}}}\, q^{N_{\mathrm{c}}}\,, \end{aligned} \tag{4.3}$$

where N_{b} is the total number of active bonds on the lattice in a configuration, N_{m} the total number of missing bonds, and N_{c} the total number of clusters. The sum is over all configurations of active bonds. Note that in the q-state Potts case every spin of a cluster must be in the same state of the q possible states, but the different clusters are completely uncorrelated, so the state of a cluster is independent of the states of all the other clusters. In (2.106) and (4.3) no symmetry-breaking field singling out one of the states was included, and hence integrating out the spin states for any cluster simply gives a factor q in (4.3).

How can one use (4.3) to construct a Monte Carlo algorithm? The recipe proposed by *Swendsen* and *Wang* [4.1, 4.46] proceeds in two steps: first, the spin configuration of the lattice is used to assign to the system a configuration of active bonds, using (4.2) and attaching a bond if p exceeds the random number ζ drawn uniformly from the interval [0,1] and attaching no bond if $p < \zeta$. As mentioned above, no bonds are assigned between sites with different spin values. This step yields a configuration of bonds that form clusters of spins, with all spins in each cluster having the same value.

The second step consists in choosing at random new values of the spins, with the constraint that every spin in a cluster must have the same new value. In this way, a new spin configuration of the lattice will result, and the process is repeated again and again.

This algorithm is ergodic, since the probability of going from a configuration to any other configuration in a single "sweep" through the lattice is nonzero, and it satisfies detailed balance.

We do not discuss the extensions to include magnetic fields, antiferromagnetic interactions [4.46], isotropic spins or lattice gauge models [4.47], etc., but rather treat the analysis of results in the framework of simulations utilizing this algorithm [4.25].

For expressing the variables of interest for an Ising model ($q = 2$) in terms of clusters, it is convenient to use two "cluster coordinates" [4.105], the cluster magnetization $m_{\mathrm{cl}} = \pm l$ for a cluster containing l sites (the sign specifies the orientation of the spins inside the cluster; we label clusters by the index "cl"), and the number of active bonds in the cluster which we denote as $u_{\mathrm{cl}} p$. Defining then the number of clusters with these properties per lattice site as $p(m_{\mathrm{cl}}, u_{\mathrm{cl}})$, magnetization and energy per spin for a lattice of coordination number z are given by (remember that $H = 0$)

$$\langle M \rangle = \sum_{m_{\mathrm{cl}}} m_{\mathrm{cl}} P(m_{\mathrm{cl}}), \quad P(m_{\mathrm{cl}}) \equiv \sum_{u_{\mathrm{cl}}} p(m_{\mathrm{cl}}, u_{\mathrm{cl}}), \tag{4.4}$$

$$\begin{aligned} E &= \langle \mathcal{H}_{\mathrm{Ising}} \rangle / N = -\frac{z}{2} J \left(\sum_{m_{\mathrm{cl}}} \sum_{u_{\mathrm{cl}}} p(m_{\mathrm{cl}}, u_{\mathrm{cl}}) - 1 \right) \\ &= -\frac{z}{2} J (p \langle N_{\mathrm{b}} \rangle / N - 1), \end{aligned} \tag{4.5}$$

recalling that N_{b} is the total number of active bonds in a configuration, and N is the number of lattice sites ($N = L^d$ for a hypercubic lattice of linear dimension L in d dimensions). In addition, the specific heat and susceptibility can be expressed in terms of clusters, e.g.,

$$\begin{aligned} C &= \partial E / \partial T = \left[1 / \left(N k_{\mathrm{B}} T^2 \right) \right] \left(\langle \mathcal{H}_{\mathrm{Ising}}^2 \rangle - \langle \mathcal{H}_{\mathrm{Ising}} \rangle^2 \right) \\ &= \frac{1}{4} z^2 J^2 / \left(N k_{\mathrm{B}} T^2 p^2 \right) \left[\langle N_{\mathrm{b}}^2 \rangle - \langle N_{\mathrm{b}} \rangle^2 - (1 - p) \langle N_{\mathrm{p}} \rangle \right]. \end{aligned} \tag{4.6}$$

Splitting off the contribution of the largest cluster in the system, which we denote as m_{cl}^{∞}, from $P(m_{\mathrm{cl}})$,

$$P(m_{\mathrm{cl}}) \equiv P'(m_{\mathrm{cl}}) + (1/N) \delta_{m_{\mathrm{cl}}, m_{\mathrm{cl}}^{\infty}}, \tag{4.7}$$

we see that the absolute value of magnetization differs from the percolation probability $\langle P_{\infty} \rangle$, due to contributions from the smaller clusters

$$\langle |M| \rangle = \left\langle \left| \frac{m_{\mathrm{cl}}^{\infty}}{N} + \sum_{m_{\mathrm{cl}}} m_{\mathrm{cl}} P'(m_{\mathrm{cl}}) \right| \right\rangle, \quad \langle P_{\infty} \rangle = \left\langle \left| m_{\mathrm{cl}}^{\infty} \right| \right\rangle / N. \tag{4.8}$$

Now the susceptibility for $T > T_c$ is just the analog of the percolation susceptibility, (2.60), namely

$$k_B T \chi = k_B T (\partial \langle M \rangle / \partial H)_{T, H=0} = \sum_{m_{cl}} m_{cl}^2 P(m_{cl}) = \sum_l l^2 n_l \tag{4.9}$$

since $P(m_{cl}) + P(-m_{cl}) = n_l$, the number of clusters of size l, apart from a different normalization factor $(1/p)$ used there, and the fact that one removes the largest cluster from the summation in (2.60) in order to work with the same formula on both sides of the percolation threshold. In (4.9), the largest cluster should not be omitted if one wishes to maintain the fluctuation relation $k_B T \chi = N \langle M^2 \rangle$ for $T \geq T_c$ and $H = 0$; see (2.77a).

Now for $T < T_c$ one must single out the contribution from the largest cluster (that becomes the percolation cluster for $N \to \infty$ and then carries the spontaneous magnetization) to obtain [4.25]

$$\begin{aligned} k_B T \chi' \equiv N \left(\langle M^2 \rangle - \langle |M| \rangle^2 \right) &= \sum_l{}' l^2 n_l + N \left(\langle P_\infty^2 \rangle - \langle |M| \rangle^2 \right) \\ &\approx \sum_l{}' l^2 n_l + N \left(\langle P_\infty^2 \rangle - \langle P_\infty \rangle^2 \right). \end{aligned} \tag{4.10}$$

The obvious physical interpretation of (4.10) is, of course, that the response function for $T < T_c$ picks up contributions both from all finite clusters (the term $\sum_l{}' l^2 n_l$, only considered in the percolation problem) and from the fluctuations in size of the largest (percolating) cluster.

It turns out that estimating χ, χ' from these relations in terms of clusters is advantageous in comparison with the standard magnetization fluctuation relations: Equations (4.9) and (4.10) already exploit the fact that there are no correlations between different clusters, thus the statistical noise is reduced. The right-hand sides of these equations hence are examples for the use of "improved estimators". Figure 4.1 shows an example for the $d = 2$ Ising square lattice. It is clearly seen that for finite systems the percolation probability $\langle P_\infty \rangle$ is always smaller than $\langle |M| \rangle$, as expected from (4.8), although in the limit $N \to \infty$ both quantities converge to the spontaneous magnetization. Note, however, that even for $N \to \infty$ the term $N(\langle P_\infty^2 \rangle - \langle P_\infty \rangle^2)$ must not be neglected in $k_B T \chi'$ in comparison to $\sum_l{}' l^2 n_l$ for $T < T_c$, although it is negligible for $T > T_c$. This observation corroborates our conclusion; cf. (2.80d) and Fig. 2.15, that due to spontaneous symmetry breaking one needs to use different fluctuation formula above and below T_c, unlike the percolation case where only $\sum_l{}' l^2 n_l$ applies on both sides of the percolation threshold.

As an example of application for which the use of a cluster algorithm was indispensable, we mention the study of shape effects on finite-size scaling [4.106]. Let us consider anisotropic subsystems of shape $L_\perp / L_\parallel$ with the linear dimensions $L_\parallel$, $L_\perp$ in the x, y directions different from each other, for a two-dimensional Ising system (the total system size $L \times L \to \infty$). At T_c from the exact results on the critical correlations ($g(r) \propto r^{-\eta}$, $\eta = \frac{1}{4}$), one can deduce that [4.106]

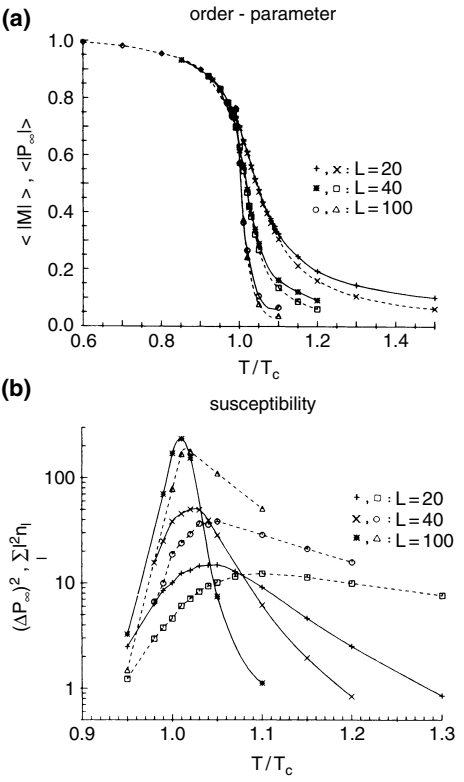


Fig. 4.1 **a** Magnetization (*solid curves*) and percolation probability (*broken curves*) for the $d = 2$ nearest-neighbor Ising ferromagnet plotted versus reduced temperature for the three system sizes indicated. Periodic boundary conditions were used throughout, and all data were generated with the algorithm of *Swendsen* and *Wang* [4.1]. (From [4.25]). **b** Normalized fluctuation of the largest cluster, $N(\langle P_\infty^2\rangle - \langle P_\infty\rangle^2)$ (*solid curves*) and second moment of the cluster size distribution, $\sum' l^2 n_l$ (*broken curves*) plotted versus T/T_c, for the same model as in **a**. (From [4.25])

$$k_B T_c \chi(T_c) = (L_\parallel L_\perp)^{1-\eta/2} \tilde{\chi}(L_\parallel / L_\perp), \quad \tilde{\chi}(\zeta) = \tilde{\chi}\left(\frac{1}{\zeta}\right) \propto \zeta^{-\eta/2}. \tag{4.11}$$

This structure can also be deduced from conformal invariance [4.107]. Since [4.104] $\gamma/\nu = 2 - \eta$, extracting a factor $(L_\parallel/L_\perp)^{1-\eta/2}$ from the scaling function $\tilde{\chi}$, (4.11) can be rewritten in the form $k_B T_c \chi(T_c) = L_\parallel^{\gamma/\nu} \overset{\approx}{\chi}(L_\parallel/L_\perp)$, with $\overset{\approx}{\chi}$ another scaling function. This behavior is tested in Fig. 4.2, using subsystems in the range $4 \le L_\parallel$, $L_\perp \le 128$, for $L = 1{,}024$ (Fig. 4.2). Of course, $L \gg L_\parallel$, $L_\perp$ is indispensable for (4.11) to be valid. Since with a standard single spin-flip algorithm the relaxation time scales as $\tau \propto L^z$ with $z \approx 2.14 \pm 0.05$ [4.108], it would be extremely hard to generate many uncorrelated configurations of a lattice of this size $L = 1{,}024$ at T_c. Thus the present problem can only be treated with the help of a cluster algorithm.

One caveat, however, must be mentioned: cluster algorithms seem to be much more sensitive to correlations among the pseudorandom numbers than the single spin-flip type algorithms [4.109]. For this reason it is often advisable to randomly mix in single spin-flip sweeps through the lattice, in addition to the "cluster moves" [4.109].

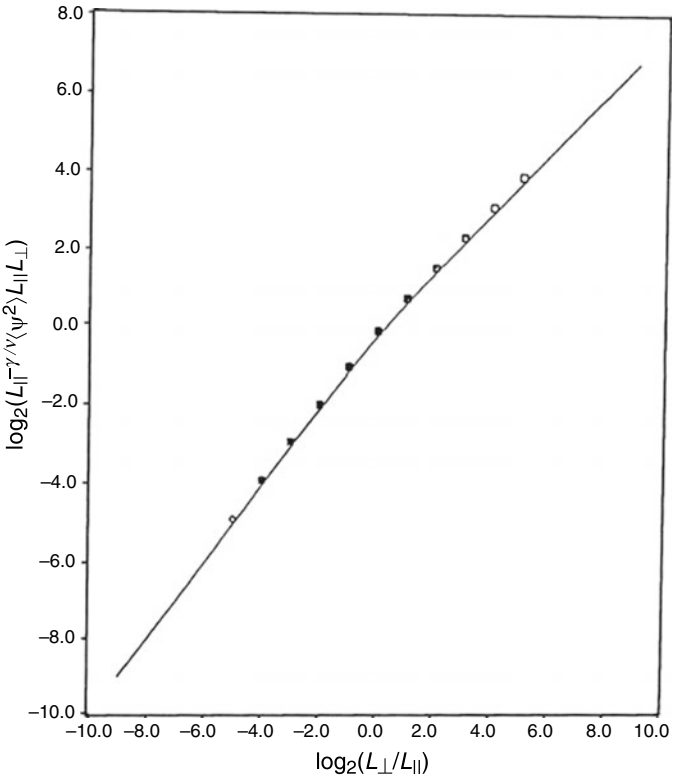


Fig. 4.2 Log_2–$\log_2$ plot of $k_\mathrm{B} T_\mathrm{c} \chi(T_\mathrm{c}) L_\parallel^{-\gamma/\nu}$ versus $L_\perp / L_\parallel$ using subsystems of the linear dimensions $L_\perp$, $L_\parallel = 4, 8, 16, 32, 64$ and 128, for a $1{,}024 \times 1{,}024$ square lattice studied with the *Swendsen–Wang* algorithm [4.1]. The curve is an approximate calculation of the scaling function. (From [4.106])

4.3 Reweighting Methods in the Study of Phase Diagrams, First-Order Phase Transitions, and Interfacial Tensions

In this section, we return to the use of "histogram reweighting techniques" already alluded to in (4.1). We first treat the single histogram method and consider over which region $\Delta\beta$ of the inverse temperature β around the reference state at β_o a reweighting is possible [4.63–4.65]. First we note that the energy distribution $P(E, T_\mathrm{o})$ has a peak of width, cf. (2.33)

$$\beta_\mathrm{o} \delta E = \left[C(\beta_\mathrm{o}) k_\mathrm{B} L^{-d} \right]^{1/2}, \quad T \neq T_\mathrm{c}, \tag{4.12}$$

$C(\beta)$ being the specific heat at inverse temperature β. Also the shift of the position of the maximum of $P(E, T)$ relative to the maximum of $P(E, T_\mathrm{o})$ is controlled by the specific heat, $\Delta E = C(\beta_\mathrm{o})(T - T_\mathrm{o}) \approx C(\beta_\mathrm{o})(\beta_\mathrm{o} - \beta)/\beta_\mathrm{o}^2$. The maximum

$\Delta\beta = \beta_o - \beta$ that gives reliable results occurs when $\Delta E \propto \delta E$, so that

$$(\Delta\beta)_{\max}/\beta_o \propto \left[C(\beta_o)L^d/k_B\right]^{-1/2}, \quad T \neq T_c, \tag{4.13}$$

i.e., the range of $\Delta\beta$ over which reweighting is possible decreases like the square root of the simulation volume. If we work at the critical point, however, and consider a system whose specific heat diverges like $C \propto (T - T_c)^{-\alpha}$, we must take into account finite-size rounding of this divergence to use $C(\beta_o) \propto L^{\alpha/\nu}$ if $\beta_o = 1/T_c$. In this case $(\Delta\beta)_{\max}/\beta_o \propto L^{-(d+\alpha/\nu)/2} \propto L^{-1/\nu}$, where in the last step the hyperscaling relation $d\nu = 2 - \alpha$ [4.96] was used. However, the region $\Delta\beta L^{1/\nu} = \text{const}$ is exactly the region of interest for finite-size scaling (see Chap. 2). As a consequence, a single histogram is enough to cover a finite fraction of the region needed to study the finite-size scaling behavior, and a finite number of neighboring histograms thus suffices to explore the full region of interest for finite size scaling.

Another obvious extension is that one may wish to carry out a reweighting not only with respect to temperature but also with respect to other thermodynamic variables. For example, consider a symmetrical polymer mixture [4.110–4.113]: We consider a lattice containing two types of polymers (A,B) with chain lengths $N_A = N_B = N$, the numbers of chains being n_A, n_B, with the total number $n = n_A + n_B$ being held fixed so the total polymer density is constant. The polymer chains may be modelled either as self-avoiding walks [4.113] (Sect. 2.1.3) or by the so-called bond-fluctuation model [4.114], where each effective monomer takes all 8 sites at the corners of an elementary cube of the simple cubic lattice. In any case, one assumes short-range pairwise interactions $\varepsilon_{AA} = \varepsilon_{BB}$, ε_{AB} between the effective monomers, which may lead to phase separation if $\varepsilon = \varepsilon_{AB} - (\varepsilon_{AA} + \varepsilon_{BB})/2 > 0$. The order-parameter of such an unmixing transition can then be defined as $m = (n_A - n_B)/(n_A + n_B)$. In this problem it is clearly advantageous not to work in the canonical ensemble of the mixture (all particle numbers $n_A N, n_B N$ being fixed, and hence also $m = \text{const}$) but in the semi-grand canonical ensemble of the mixture, where T and the chemical potential difference $\Delta\mu$ of the mixture are the given independent thermodynamic variables. Note that if we considered a mixture of monomers ($N = 1$) in the framework of the Ising model, $\Delta\mu$ simply corresponds to the magnetic field H, and m to the magnetization [4.115]. The semi-grandcanonical partition function $Z_{SG}(T, \Delta\mu)$ then can be related to the density of states $\Gamma(E, m)$ as

$$Z_{SG}(T, \Delta\mu) = \int_{-1}^{+1} \mathrm{d}m \int \mathrm{d}E \exp(-E/k_B T) \exp[N\Delta\mu mn/(2k_B T)] \times \Gamma(E, m)\,. \tag{4.14}$$

Note that chemical potentials μ_A, μ_B of the two species were normalized per monomer, so the Boltzmann factor $\exp[(\mu_A n_A N + \mu_B n_B N)/k_B T] = \exp[N(\mu_A + \mu_B)n/(2k_B T)] \exp[N\Delta\mu mn/(2k_B T)]$, and the first factor $\exp[N(\mu_A + \mu_B)n/(2k_B T)]$ is omitted since n is constant, and this factor hence cancels out from all averages. The Monte Carlo sampling yields a number $\mathcal{N}$ of configurations that are

distributed proportional to $P_{T,\Delta\mu}(E,m)$,

$$P_{T,\Delta\mu}(E,m) = \frac{1}{Z_{\mathrm{SG}}(T,\Delta\mu)} \exp(-E/k_{\mathrm{B}}T) \exp[N\Delta\mu mn/(2k_{\mathrm{B}}T)]\Gamma(E,m). \tag{4.15}$$

We record a "histogram" $H_{T,\Delta\mu}(E,m)$ in a simulation by just counting how often one observes the possible values of E and m (for very large lattice sizes, and generally for continuous degrees of freedom, some binning in suitable intervals δE, δm may be required, but this is not discussed further here). This histogram is now used for a whole range of neighboring values T', $\Delta\mu'$ around T, $\Delta\mu$ by reweighting

$$P_{T',\Delta\mu'}(E,m) \approx \frac{Z_{\mathrm{SG}}(T,\Delta\mu)}{Z_{\mathrm{SG}}(T',\Delta\mu')} \exp\left(\frac{\Delta\mu' mNn}{2k_{\mathrm{B}}T'} - \frac{\Delta\mu mNn}{2k_{\mathrm{B}}T}\right) \times \exp\left(\frac{E}{k_{\mathrm{B}}T} - \frac{E}{k_{\mathrm{B}}T'}\right) \frac{\mathcal{H}_{T,\Delta\mu}(E,m)}{\mathcal{N}}. \tag{4.16}$$

Figure 4.3 shows that this "single histogram extrapolation" is indeed of practical usefulness for the Monte Carlo study of the critical region of polymer mixtures [4.112]. Note that phase coexistence for a symmetric mixture occurs for $\Delta\mu = 0$ as studied there, and thus $P(m)$ is symmetric around $m = 0$, with two peaks for $T < T_{\mathrm{c}}$ (the peak for $m < 0$ corresponds to the B-rich phase, the peak for $m > 0$ corresponds to the A-rich phase) which merge into a single peak as the temperature is raised above T_{c}. Of course, Fig. 4.3 is just an illustration of the behavior postulated in (2.73) and (2.74).

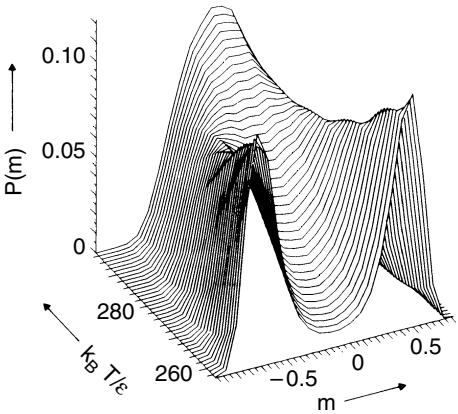


Fig. 4.3 Distribution function $P(m) = \int \mathrm{d}E\, P_{T,\Delta\mu}(E,m)$ of the order-parameter m of a polymer mixture plotted for $\Delta\mu = 0$ over a range of temperatures for $N = 128$, $-\varepsilon_{\mathrm{AA}} = -\varepsilon_{\mathrm{BB}} = \varepsilon_{\mathrm{AB}} = \varepsilon/2$, and a density $\phi = 0.5$ of occupied lattice sites, and a lattice size $L = 80$ lattice spacings. Using the bond-fluctuation model on the simple cubic lattice, $\mathcal{N} = 16{,}800$ statistically independent samples at $k_{\mathrm{B}}T/\varepsilon = 266.4$ were used to generate $P(m)$ over a range of temperatures from single-histogram extrapolation. (From [4.112])

As mentioned above, it is often desirable to combine histograms from several simulation runs to get extrapolations over a wider parameter range and to increase the accuracy [4.66]. If k simulations are performed at parameters $[(T_i, \Delta\mu_i), i = 1, \ldots, k]$, the generalization of (4.14)–(4.16) reads [$\mathcal{N}_i$ is now the number of states recorded in simulation i for the histogram $H_{T_i,\Delta\mu_i}(E, m)$],

$$\Gamma(E, m) = \sum_{i=1}^{k} w_i(E, m) \frac{1}{\mathcal{N}_i} Z_{\mathrm{SG}}(T_i, \Delta\mu_i) \exp[-N\Delta\mu_i mn/(2k_{\mathrm{B}}T)] \times \exp(E/k_{\mathrm{B}}T_i) H_{T_i,\Delta\mu_i}(E, m). \tag{4.17}$$

The weight w_i with which the histogram of the ith simulation enters is determined by minimizing the statistical error of $\Gamma(E, m)$ and is found as [4.46, 4.66, 4.111]

$$w_i(E, m) = w_i'(E, m) / \sum_{j=1}^{k} w_j'(E, m), \tag{4.18}$$

where

$$w_i'(E, m) = (1 + 2\tau_i)^{-1} \mathcal{N}_i \, Z_{\mathrm{SG}}^{-1}(T_i, \Delta\mu_i) \exp[N\Delta\mu_i mN/(2k_{\mathrm{B}}T_i)] \times \exp(-E/k_{\mathrm{B}}T_i), \tag{4.19}$$

τ_i being the correlation time (measured in Monte Carlo steps per monomer if every step is used for the recording of the histograms, while $\tau_i = 0$ if enough steps are omitted between subsequent recordings such that all $\mathcal{N}_i$ configurations that are used are statistically independent). Note that only the product $w_i Z_{\mathrm{SG}}(T_i, \Delta\mu_i)$ enters (4.17) and from (4.19) we see that no knowledge of the absolute free energy is required. Figure 4.4 gives an example where this technique was used to obtain moments $\langle |m| \rangle$, $\langle m^2 \rangle$ over a broad temperature range for three lattice sizes [4.112]. Obviously, such techniques are a great advantage for a precise estimation of the critical point: while in the traditional method (as described in Chap. 2, e.g., Fig. 2.14a) the data are generated point by point and curves drawn to guide the eye always are somewhat subjective, from the density of states based on several histograms, (4.15), the smooth curves in Fig. 4.4 result in a well-defined way, and many application examples meanwhile testify for the success and accuracy of these methods.

The full power of these histogram extrapolation techniques becomes apparent, however, when we consider asymmetric mixtures. In a lattice model containing A,B, and vacancies, the simplest type of asymmetry occurs if $\varepsilon_{\mathrm{AA}} \neq \varepsilon_{\mathrm{BB}}$: then the symmetry of the problem against interchange of A and B is broken, and phase coexistence (including criticality) no longer occurs for $\Delta\mu = 0$ but along a nontrivial curve $\Delta\mu_{\mathrm{coex}}(T)$ in the $(\Delta\mu, T)$ space. And for finding the critical point $\Delta\mu_{\mathrm{c}} = \Delta\mu_{\mathrm{coex}}(T_{\mathrm{c}})$ it is obvious that a search in the two-dimensional space point by point would be very cumbersome, while histogram extrapolations still are convenient [4.116, 4.117].

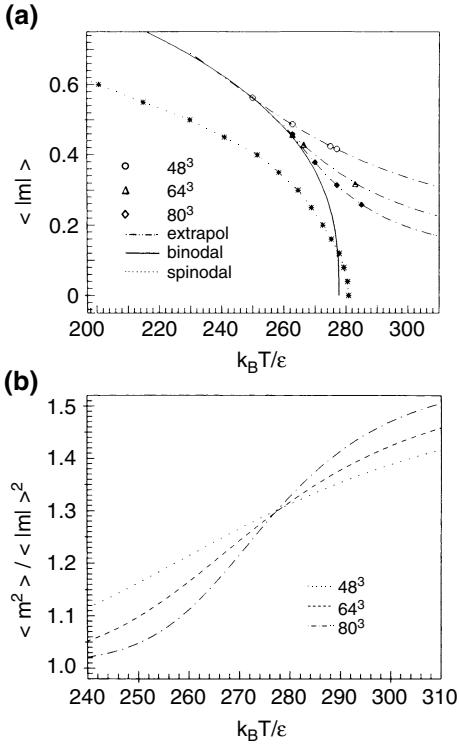


Fig. 4.4 Order-parameter $\langle|m|\rangle$ (**a**) and ratio of moments $\langle m^2\rangle/\langle|m|\rangle^2$ (**b**) plotted versus temperature, for the same model, as described in Fig. 4.3, but using three lattice sizes and multiple histogram extrapolation. The *open symbols* in part (**a**) (*triangles*, *diamonds*, and *circles*) show the data directly observed at those temperatures where actual simulations were made, while the *dash-dotted curves* result from the histogram extrapolation, as well as the curves in part (**b**), where $T_c = 277.7\varepsilon/k_B$ was extracted from the intersection of the curves. Using this estimate for T_c in a finite-size scaling analysis of the data in part (**a**), the estimate for the order-parameter (*solid curve*, labelled as "binodal") is obtained. The *dotted curve with asterisks* (marked "spinodal") illustrates the result of a mean-field concept, namely the inverse "susceptibility" $\chi^{-1}(T, m)$ is extrapolated and the spinodal curve is the locus of the points where this extrapolation vanishes, $\chi^{-1}(T, m = m_{sp}(T)) = 0$. (From [4.112])

In this case one proceeds again by the "equal weight rule", which we have already encountered in Sect. 2.3.7 as a criterion to locate a first-order phase transition: there we have considered the Ising model for $T < T_c$ as a function of the magnetic field H, and found that the two Gaussian peaks centered at positive and negative magnetization change their weights according to the Boltzmann factors $\exp(\pm m H L^d/k_B T)$. In this case symmetry requires the weights to be equal for $H = 0$, while in the present case H corresponds to $\mu' \equiv [\Delta\mu - \Delta\mu_{coex}(T)]/\varepsilon$, and finding the state of equal weight ($\mu' = 0$) is nontrivial. We can define these weights for the A-rich and B-rich phases as

$$P_{\text{A-rich}} = \int_{m^\star}^{1} P(m)\mathrm{d}m, \quad P_{\text{B-rich}} = \int_{-1}^{m^\star} P(m)\mathrm{d}m, \tag{4.20}$$

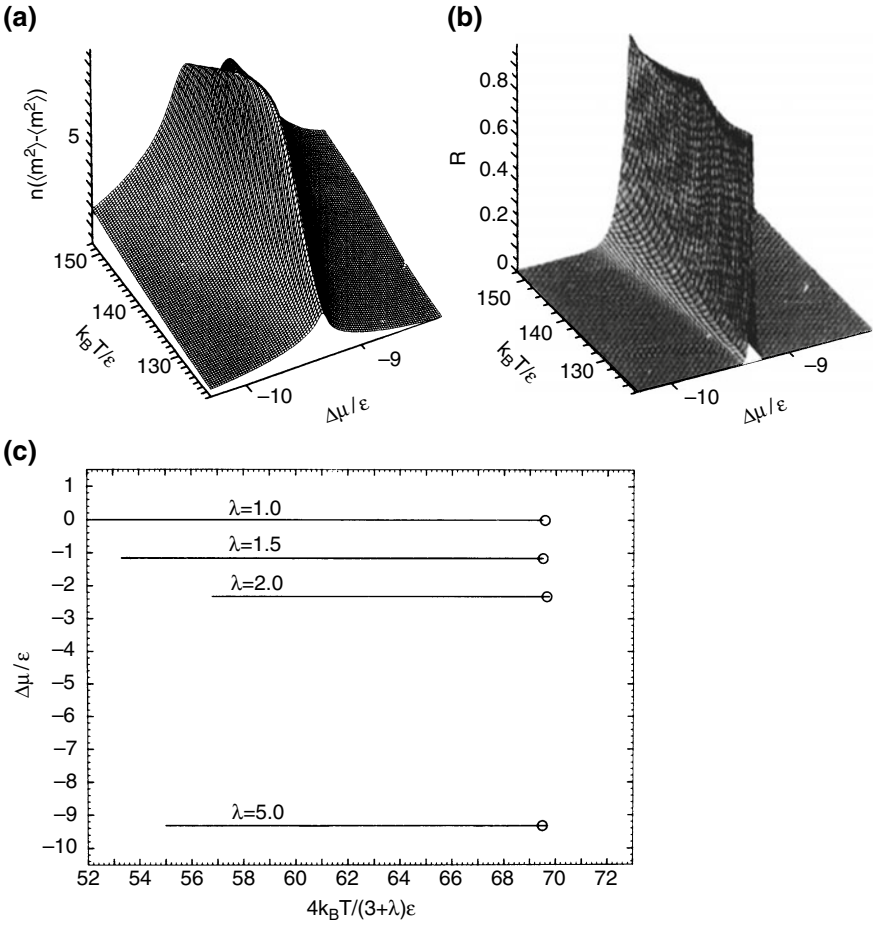


Fig. 4.5 **a** The fluctuation of the order-parameter m, $n(\langle m^2\rangle - \langle|m|\rangle^2)$ for the bond-fluctuation model of a polymer mixture with $N_A = N_B = N = 32$ but energetic asymmetry ($\varepsilon_{AA} = \lambda\varepsilon_{BB}$ with $\lambda = 5$, while $\varepsilon_{AB} = -\varepsilon_{BB} \equiv \varepsilon$). For a lattice linear dimension $L = 56$ the system contained $n = n_A + n_B = 343$ polymer chains. From a few state points for which actual simulations were performed, the fluctuation is obtained in a broad region of the $(T, \Delta\mu)$ plane by histogram reweighting. The maximum of the ridge yields a rough first estimate of the critical temperature, T_c. The order-parameter $\langle m\rangle$ along the ridge can be identified with $m^\star$ in (4.18). **b** The ratio R of the weights P_{A-rich}, P_{B-rich} plotted in the same region of the $(T, \Delta\mu)$ plane as part (**a**). Here R is defined as the minimum value of P_{A-rich}/P_{B-rich} and its inverse, so it can only vary between zero and one, and the location of the maximum in the $(T, \Delta\mu)$ plane then yields the coexistence curve $\Delta\mu = \Delta\mu_{coex}(T)$, see part (**c**). **c** Phase diagrams of the asymmetric polymer mixture in the $(T, \Delta\mu)$ plane plotted for four values of χ (for $\lambda = 1$ the mixture is symmetric, then $\Delta\mu_{coex} \equiv 0$ by symmetry, while in all other cases a nontrivial offset occurs). The critical points (marked by circles) are found by studying the ratio $\langle m^2\rangle/\langle|m|\rangle^2$ versus T along the curve $\mu_{coex}(T)$ for the three lattice sizes $L = 32$, 40 or 56, respectively. (From [4.117])

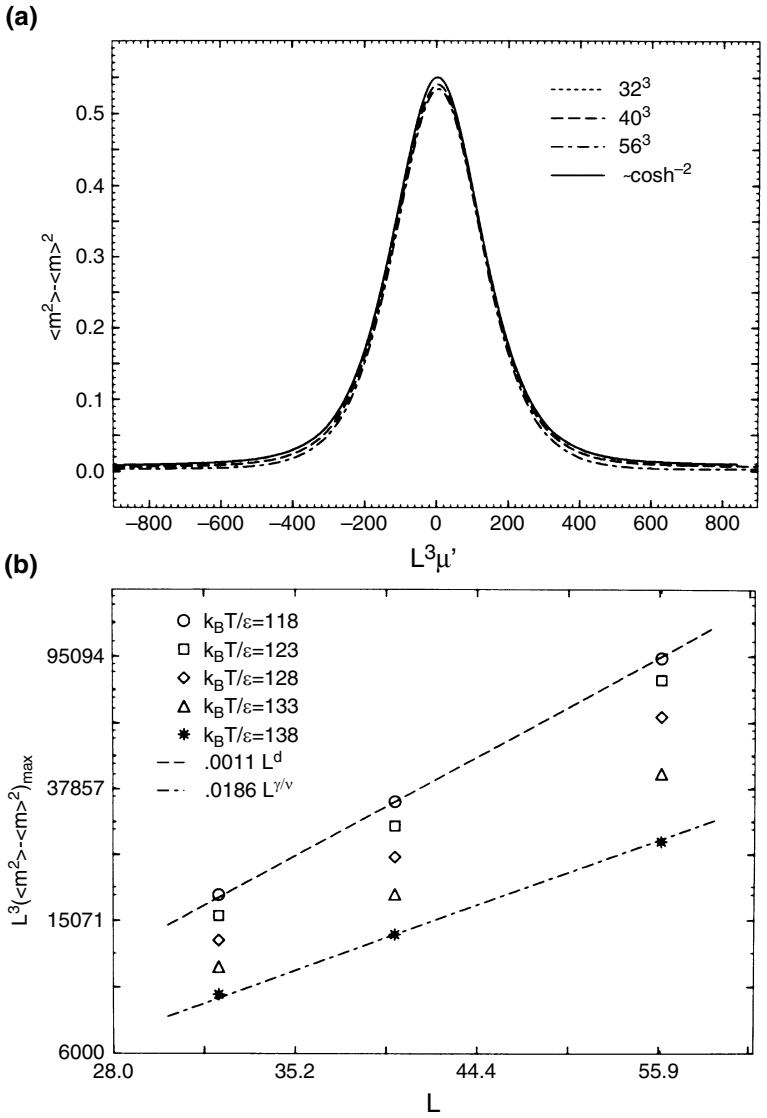


Fig. 4.6 **a** Scaled susceptibility $\langle m^2\rangle - \langle m\rangle^2$ at the first-order transition for the same model as in Figs. 4.4 and 4.5, for $k_B T/\varepsilon = 118$. Here $\mu' \equiv (\Delta\mu - \Delta\mu_{\rm coex})/32\varepsilon$. The *solid curve* shows the function $m_o^2/[\cosh(\mu' m_o L^3/k_B T)]^2$, where m_o is the value $m_{\rm coex} - m^\star$, cf. (2.97). **b** Log–log plot of the maximum value of the susceptibility $L^3(\langle m^2\rangle - \langle m\rangle^2)_{\rm max}$ versus system size, for the same model as in Figs. 4.4, 4.5, and five different temperatures. The highest temperature (*stars*) is the critical temperature, where scaling $\propto L^{\gamma/\nu}$ with $\gamma/\nu = 1.96$ is confirmed. (From [4.117])

where $m^\star$ is the location of the minimum between the two peaks of $P(m)$ {note that $m^\star = 0$ in the symmetric case, see Fig. 4.3; in the asymmetric case, estimates for $m^\star$ can conveniently be extracted from the maximum of $\langle m^2\rangle - \langle m\rangle^2$ or the reduced connected fourth order cumulant $\langle m^4\rangle_{\rm conn}/[3\langle m^2\rangle^2]$, as shown by *Deutsch* [4.117]}. Figure 4.5 demonstrates that indeed very good accuracy can be obtained. The fact that from Monte Carlo runs just at a few state points $\{T_i,\ \Delta\mu_i\}$ one is able to obtain all thermodynamic functions of interest in a broad region of the $(T, \Delta\mu)$ plane is absolutely crucial for this analysis. The conclusions about this phase diagram (Fig. 4.5c) are corroborated by the finite-size scaling analysis of the first-order phase transition as well: Figure 4.6 gives a three-dimensional counterpart to Fig. 2.21b, demonstrating that the concepts developed in Chap. 2 for the simple Ising model do straightforwardly carry over to much more complex systems, such as asymmetric polymer mixtures.

Finally, we mention that sampling of the minimum of distributions as shown in Fig. 4.3 yields information on the interfacial tension [2.70, 4.70–4.75].

|

4.4 Some Comments on Advances with Finite-Size Scaling Analyses

The extension of finite-size scaling techniques to the case of mixtures with energy asymmetry [4.116, 4.117] is rather straightforward, since in the $(\Delta\mu, T)$ plane the coexistence curve $\mu_{\rm coex}(T)$ still has a (nearly) vanishing slope at criticality, $r = (T_{\rm c}/\varepsilon)d\mu_{\rm coex}(T)/dT|_{T_{\rm c}} \approx 0$ (Fig. 4.5c). The situation is no longer so simple, when r is appreciably different from zero, as happens for polymer mixtures with asymmetry in the chain lengths [4.118], polymer solutions [4.90] and ordinary Lennard–Jones fluids [4.85–4.88, 4.93, 4.94]. Then, the scaling fields and scaling operators are no longer $t = \varepsilon/k_{\rm B}T_{\rm c} - \varepsilon/k_{\rm B}T$ (temperature-like variable), $\mu' = (\Delta\mu - \Delta\mu_{\rm crit})/k_{\rm B}T_{\rm c}$ (field-like variable), ϱ (concentration or density, respectively), u (energy density), but linear combinations thereof [4.119]

$$\tau = t + s\mu', \qquad h = \mu' + rt, \tag{4.21}$$

and thermodynamically conjugate to these fields (τ, h) are the order-parameter $\mathcal{M}$ and energy-like density $\mathcal{E}$,

$$\mathcal{M} = (\varrho - su)/(1 - sr), \quad \mathcal{E} = (u - r\varrho)/(1 - sr). \tag{4.22}$$

The geometric meaning of the other field-mixing parameter s is given as $s = \tan\psi$ where ψ is the angle between h and the ordinate at criticality. In an Ising ferromagnet, of course, this problem is completely absent, because the order-parameter (magnetization) is odd in the spin variables, the exchange energy is even in the spin

variables, and already this symmetry implies that there cannot be any field mixing. Already for the liquid gas transition the situation is completely different, of course, since there is no simple symmetry relating liquid and gas, and therefore the density ϱ is not the optimal choice for the order-parameter field. Subtracting off the scaling operators at criticality,

$$\delta\mathcal{M} \equiv \mathcal{M} - \langle\mathcal{M}\rangle_{\rm crit}, \quad \delta\mathcal{E} \equiv \mathcal{E} - \langle\mathcal{E}\rangle_{\rm crit}, \tag{4.23}$$

the joint distribution $p_L(\mathcal{M},\mathcal{E})$ then acquires a simple finite size scaling form [4.93, 4.94]

$$P_L(\mathcal{M},\mathcal{E}) \approx (a_{\mathcal{M}}a_{\mathcal{E}})^{-1} L^{d+\beta/\nu-1/\nu}\, \tilde{p}\big(a_{\mathcal{M}}^{-1}L^{\beta/\nu}\delta\mathcal{M}, a_{\mathcal{E}}^{-1}L^{d-1/\nu}\delta\mathcal{E}, a_{\mathcal{M}}L^{d-\beta/\nu}h, a_{\mathcal{E}}L^{1/\nu}\tau\big), \tag{4.24}$$

where β and ν are critical exponents of the order-parameter and the correlation length as usual, $a_{\mathcal{M}}, a_{\mathcal{E}}$ are two (non-universal, i.e., system-dependent) scale factors, while the scaling function $\tilde{p}$ then is universal. From (4.22)–(4.24) we recognize that integrating out $\mathcal{E}$ the scaling of $P_L(\mathcal{M}) = \int d\mathcal{E} p_L(\mathcal{M},\mathcal{E})$ is exactly the same as the scaling of the order-parameter distribution $p_L(s)$ considered in (2.79). In contrast, if we consider the simple distribution $p_L(\phi,u)$ rather than $p_L(\mathcal{M},\mathcal{E})$, which can be found by using (4.22) in (4.24), we see that scaling powers $L^{d-1/\nu}(\varrho-\varrho_{\rm crit})$ and $L^{\beta/\nu}(u-u_{\rm crit})$ would also result. As a consequence, the scaling behavior of ϱ and u is not simple, and some well-known recipes (like extracting the specific heat exponent α from the energy fluctuations, which works in an Ising model as $C^{\rm Ising}_{H=0} = L^d(\langle u^2\rangle - \langle u\rangle^2)/k_{\rm B}T^2 \propto L^{\alpha/\nu}$ are no longer true: energy density fluctuations in asymmetric systems also scale with the susceptibility exponent γ

$$L^d\left(\langle u^2\rangle - \langle u\rangle^2\right)/k_{\rm B}T^2 \propto L^{\gamma/\nu}, \tag{4.25}$$

while the specific heat exponent α shows up in the fluctuation of the energy-like variable [4.93, 4.94]

$$L^d\left(\langle \mathcal{E}^2\rangle - \langle\mathcal{E}\rangle^2\right)/k_{\rm B}T^2 \propto L^{\alpha/\nu}. \tag{4.26}$$

Another consequence of (4.24) is that the scaling function $\tilde{P}$ of $P_L(\mathcal{M})$ is identical to the scaling function of the Ising model, considered in (2.79). The latter can be recorded with fairly good accuracy, however, and hence *Wilding* [4.93, 4.94] suggested the use of this information for the analysis of the asymmetric systems, by adjusting the critical parameters $\{\varepsilon/k_{\rm B}T_{\rm c}, \mu/k_{\rm B}T_{\rm c}\}$ such that an optimal fit of $\tilde{P}$ onto the Ising model scaling function is achieved. In this way one can avoid the otherwise necessary variation of L over a wide range, and explore the critical region of asymmetric systems with manageable effort. While far below $T_{\rm c}$ phase coexistence of asymmetric off-lattice models of fluids can be conveniently studied with the so-called "Gibbs ensemble" [4.120, 4.121, 4.122, 4.123], near the critical point the

"field mixing" analysis techniques as developed by *Wilding* and coworkers [4.85–4.94] is the method of choice.

We conclude this section by drawing attention to the problems that finite-size scaling methods still encounter due to crossover [4.95] from one "universality class" [4.96] to another [4.92, 4.97–4.101]. There are many examples where this occurs for physical systems. (1) Heisenberg ferromagnets with weak uniaxial anisotropy show a crossover from the Heisenberg class to the Ising class as T_c is reached [4.95, 4.124]. (2) Ferromagnets weakly diluted with a non-magnetic material exhibit a crossover to a new class of critical behavior typical for randomly quenched disorder (see, e.g., [4.125] for a review). Another crossover is expected for systems in random fields [4.126]. Monte Carlo studies of systems with quenched disorder would be very valuable, since other techniques are not so conclusive about the properties of these systems, but the analysis of the simulations is manifestly hampered by crossover problems [4.127–4.131]. (3) Systems close to a multicritical point, e.g., a tricritical point [4.92, 4.132] cross over from multicritical behavior to "ordinary" critical behavior. (4) Systems with a large but finite range of interaction cross over from Landau-like critical behavior towards a nontrivial critical behavior close to T_c [4.99, 4.100, 4.133]. Analogous crossover to Landau behavior also occurs in symmetrical polymer mixtures, when the chain length $N \to \infty$ [4.97, 4.98, 4.134] and in Ising films with competing walls when their thickness $D \to \infty$ [4.101].

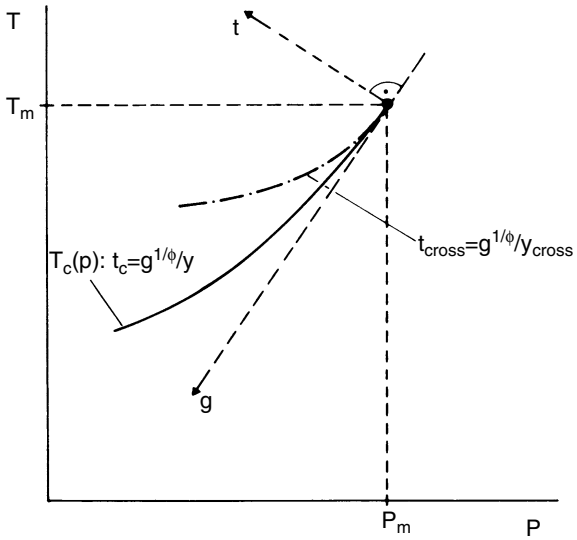


Fig. 4.7 Schematic phase diagram of a system exhibiting crossover between "ordinary" critical phenomena along the line $T_c(p)$, $p < p_m$ and the special point $p = p_m$, $T_m = T_c(p_m)$, which may be a multicritical point, for instance. Considering the approach to the critical line along an axis parallel to the t-axis, one will observe critical behavior associated with the special point, as long as one stays above the *dash-dotted curve* describing the centre of the crossover region. Only in between this dash-dotted line and the critical line (*solid curve*) the correct asymptotic behavior for $p < p_m$ can be seen. (From [4.97])

All such cases can be schematically described, as shown in Fig. 4.7: varying a parameter p the critical line $T_c(p)$ in the (T, p) plane reaches a special point $T_c(p = p_m) = T_m$ characterized by a unique critical behavior, which is distinct from the behavior along $T_c(p < p_m)$. For all $p < p_m$, the same type of critical behavior occurs but the region where it actually can be observed shrinks to zero smoothly as $p \to p_m$. Introducing scaling variables g [tangential to $T_c(p)$ in the point T_m, p_m] and t [normal to $T_c(p)$ at $p = p_m$], both the critical line and the centre of the crossover region can be expressed in terms of the crossover exponent ϕ [4.95],

$$t_c = g^{1/\phi}/y_c, \quad t_{\text{cross}} = g^{1/\phi}/y_{\text{cross}}, \tag{4.27}$$

y_c, y_{cross} being constants. More precisely, the singular part of the free energy $F^{\text{sing}}(T, H, p)$ becomes (H is the field conjugate to the order-parameter m)

$$F^{\text{sing}}(T, H, p) = t^{2-\alpha_m} \tilde{F}(Ht^{-(\beta_m+\gamma_m)}, g^{1/\phi}/t), \tag{4.28}$$

α_m, β_m, γ_m being the standard critical exponents [4.96] in the point T_m, p_m, and $\tilde{F}(x, y)$ is a scaling function. This function has at $y = y_c$ a singularity described by critical exponents α, β, γ characteristic of the universality class at the critical line, e.g.,

$$\tilde{F}(0, y) \propto (y - y_c)^{2-\alpha}, \tag{4.29}$$

while for $y \ll y_{\text{cross}} = g^{1/\phi}/t_{\text{cross}}$ the y-dependence of $\tilde{F}(x, y)$ can be neglected, $F_{\text{sing}} \propto t^{2-\alpha_m}$ for $t \gg t_{\text{cross}}$.

When we wish to study a problem such as described above by finite-size scaling, it is crucial to include the variable $g^{1/\phi}/t$ in the finite-size scaling description, since it matters whether for $t \to 0$ first the crossover (Fig. 4.7) or first the finite-size rounding (described by the variable $L^{1/\nu_m}t$, for instance) sets in (here we have assumed that L scales with the correlation length $\xi_m \propto t^{-\nu_m}$ at the special point). Then, the generalization of the finite-size scaling assumption for the order-parameter distribution, (2.79), is [4.97]

$$\begin{aligned} P(t, H, m, g, L) = & L^{\beta_m/\nu_m} \tilde{P}(L^{1/\nu_m}t, L^{(\beta_m+\gamma_m)/\nu_m}H, \\ & L^{\beta_m/\nu_m}m, L^{\phi/\nu_m}g). \end{aligned} \tag{4.30}$$

For $H = 0$ and $t = t_c$ one finds that moments and cumulants still exhibit a nontrivial behavior [4.97]

$$\langle |m|^k \rangle = L^{-k\beta/\nu} g^{(k\nu_m/\phi)[(\beta_m/\nu_m)-(\beta/\nu)]} \tilde{m}_k(L^{\phi/\nu_m}g), \tag{4.31}$$

$$U_L = 1 - \langle m^4 \rangle / [3\langle m^2 \rangle^2] = \tilde{U}(L^{\phi/\nu_m}g). \tag{4.32}$$

Only if L exceeds by far a crossover length $\xi_{\text{cross}} \propto g^{-\nu_m/\phi}$, $L \gg \xi_{\text{cross}}$, do we see the simple finite-size scaling behavior, with $\tilde{U}(\infty) = U^\star$ being the "cumulant crossing

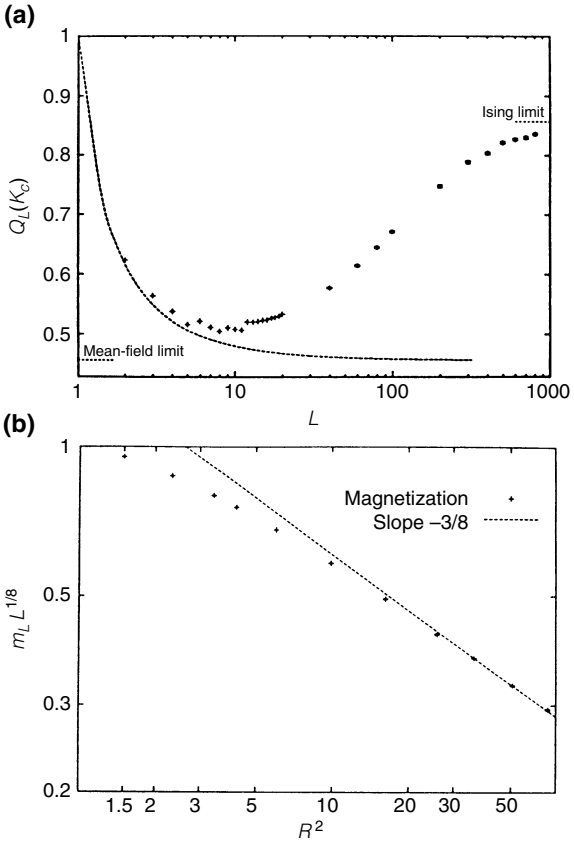


Fig. 4.8 **a** Ratio $Q_L(K_c) \equiv \langle m^2\rangle_L^2/\langle m^4\rangle_L$ for the two-dimensional Ising model where every spin interacts with all the $z = 436$ neighbors inside a maximum range of interaction $R_m = \sqrt{140}$ lattice spacings plotted at the critical coupling $K_c = J/k_B T_c = 0.0023464$ versus the system size (*discrete points*). For large L, $Q_L(K_c)$ approaches the Ising limit $Q_I = 0.856216$ [4.135], while for decreasing L, $Q_L(K_c)$ approaches the mean-field limit $Q_{MF} = 8\pi^2/[\Gamma(1/4)]^4 \approx 0.456947$ [4.136], until the system size becomes smaller than the range R_m, and other finite-size effects come into play. To illustrate that the system is indeed mean-field like for these system sizes, Q is also plotted for finite systems in which all spins interact equally strong (*dashed curve*). The points for $R_m^2 = 140$ indeed approach this curve for small L. **b** Critical amplitude of the magnetization at $K_c(R)\langle|m|\rangle_{K_c} = d_o(R)L^{-1/8}$, plotted versus R^2 on a log–log plot, where $R^2 = \sum_{j\neq i} |\boldsymbol{r}_i - \boldsymbol{r}_j|^2/z$ with $|\boldsymbol{r}_i - \boldsymbol{r}_j| \leq R_m$. *Dotted line* shows the slope predicted in (4.37). (From [4.100])

point" discussed in Sect. 2.3, and the moments are simple power laws, $\langle|m|^k\rangle \propto L^{-k\beta/\nu}$.

The above treatment is only valid if both sets of exponents $(\alpha_m, \beta_m, \gamma_m, \nu_m)$ and $(\alpha, \beta, \gamma, \nu)$ satisfy hyperscaling relations [4.96], however,

$$d = (2 - \alpha_m)/\nu_m = (2\beta_m + \gamma_m)/\nu_m = (2 - \alpha)/\nu = (2\beta + \gamma)/\nu. \tag{4.33}$$

The important case where for $d = 2$ or $d = 3$ dimensions crossover to Landau-like mean-field behavior occurs, needs separate discussion, as can be anticipated from Sect. 2.3.6. When the range of interaction R tends to infinity, we may choose $g = 1/R^d$, since then the variation of T_c with g near $g = 0$ is linear (apart from possible logarithmic corrections [4.100]), compatible with Fig. 4.7. As discussed in Sect. 2.3.6, for hypercubic geometry a simple extension of finite-size scaling is possible if the correlation length ($\xi_m \propto t^{-\nu_m} = t^{-1/2}$) is replaced by the "thermodynamic length" [2.76] $l_T \propto t^{-(\gamma_m+2\beta_m)/d} = t^{-2/d}$, cf. (2.92). Now (4.30)–(4.32) are replaced by ($2\beta_m = \gamma_m = 1$) [4.97].

$$P(t, H, m, R^{-d}, L) = L^{d/4}\tilde{P}(L^{d/2}t, L^{3d/4}H, L^{d/4}m, L^{(4-d)d/4}R^{-d}), \tag{4.34}$$

$$\langle|m|^k\rangle = L^{-dk/4}\tilde{m}_k(L^{(4-d)d/4}R^{-d}), \tag{4.35}$$

$$U_L = \tilde{U}(L^{(4-d)d/4}R^{-d}). \tag{4.36}$$

Here the crossover exponent $\phi = (4 - d)/2$ was invoked [4.97] from the Ginzburg criterion [4.133], but a more rigorous derivation using renormalization-group methods has recently been presented [4.100]. Equation (4.34) shows that L now has to be compared with a crossover length $l_{\text{cross}} \propto R^{4/(4-d)}$ ($= R^2$ in $d = 2$ dimensions). Note also the singular R-dependence of the asymptotic critical amplitude at t_c, namely

$$\langle|m|\rangle \propto L^{-\beta/\nu}R^{-(d-4\beta/\nu)/(4-d)} = L^{-1/8}R^{-3/4}(d = 2). \tag{4.37}$$

Figure 4.8 shows that these considerations indeed describe the actual behavior correctly, U_L interpolates between the theoretically expected limits [4.135, 4.136]. Note that for this study it was necessary to go to sizes $L \gg l_{\text{cross}}$ in order to be able to find the critical couplings $K_c(R)$ reliably: Thus lattice sizes up to 800×800 had to be used [4.100] which for these long-range interaction models was possible only through the availability of an efficient cluster algorithm [4.50]. Thus, this study [4.100] is a good example to demonstrate that both sophisticated new algorithms [4.50] and refinements in the finite-size scaling analysis [4.99, 4.100] are crucial for making progress. Similarly, combination of a new cluster algorithm [4.137] for the Blume–Emery–Griffiths model [4.138] with finite-size scaling promises to be useful for studying first-order transitions.

Chapter 5
Quantum Monte Carlo Simulations: An Introduction

5.1 Quantum Statistical Mechanics Versus Classical Statistical Mechanics

To be specific, let us consider for the moment the problem of N atoms in a volume V at temperature T, and we wish to calculate the average of some observable A which in quantum mechanics is described by an operator $\hat{A}$. Then the answer to this problem given by quantum statistical mechanics is

$$\langle \hat{A} \rangle = Z^{-1} \mathrm{Tr} \exp\left(-\hat{\mathcal{H}}/k_{\mathrm{B}}T\right) \hat{A} = Z^{-1} \sum_n \langle n| \exp\left(-\hat{\mathcal{H}}/k_{\mathrm{B}}T\right) \hat{A}|n\rangle, \tag{5.1}$$

where $\hat{\mathcal{H}}$ is the Hamiltonian of the system, and the trace is written here symbolically as the sum over a discrete set of states $\{|n\rangle\}$ which we assume orthonormal ($\langle n|m\rangle = \delta_{nm}$) and complete ($\sum_n |n\rangle\langle n| = \hat{1}$ where $\hat{1}$ is the identity operator). Correspondingly, the partition function Z is

$$Z = \mathrm{Tr} \exp\left(-\hat{\mathcal{H}}/k_{\mathrm{B}}T\right) = \sum_n \langle n| \exp\left(-\hat{\mathcal{H}}/k_{\mathrm{B}}T\right) |n\rangle. \tag{5.2}$$

The Hamiltonian $\hat{\mathcal{H}}$ can be written, assuming a situation where relativistic effects as well as explicit consideration of the spins of the particles can be neglected, and so the simple description in terms of the Schrödinger equation applies,

$$\hat{\mathcal{H}} = \sum_{j=1}^{N} \frac{\hat{\boldsymbol{p}}_j^2}{2m} + \sum_{i<j} \hat{V}\left(\hat{\boldsymbol{r}}_i - \hat{\boldsymbol{r}}_j\right) = \sum_{j=1}^{N} \hat{E}_j^{\mathrm{kin}} + \sum_{i<j} \hat{V}_{ij}, \tag{5.3}$$

where $\hat{\boldsymbol{p}}_i$ is the momentum operator of the ith atom (all atoms are assumed to have the mass m, and $\hat{\boldsymbol{r}}_i$ is the position operator, and we have assumed pairwise interactions between the particles described by the potential V).

K. Binder and D. W. Heermann, *Monte Carlo Simulation in Statistical Physics*,
Graduate Texts in Physics, https://doi.org/10.1007/978-3-030-10758-1_5

Now the basic reason why quantum statistical mechanics differs for this problem from classical statistical mechanics, as was assumed in the first chapter of this book, is that momentum and position operators of a particle do not commute,

$$\left[\hat{\boldsymbol{r}}_j, \hat{\boldsymbol{p}}_j\right] = i\hbar, \tag{5.4}$$

and hence also the commutator of kinetic and potential energy of a particle is non-vanishing, $\left[\hat{E}_j^{\mathrm{kin}}, \hat{V}_{ij}\right] \neq 0$. As a corollary of this statement, it can be easily seen – and this is worked out in most standard textbooks on statistical mechanics – that neglect of (5.4) for the evaluation of averages reduces the problem to classical statistical mechanics. Indeed, one can then write, using for $|n\rangle$ eigenstates of the position operators $\hat{\boldsymbol{x}}_i$ so $\sum_n$ becomes $\int \mathrm{d}\boldsymbol{x}_1 \int \mathrm{d}\boldsymbol{x}_2 \cdots \int \mathrm{d}\boldsymbol{x}_N$,

$$\begin{aligned} Z = &\int \mathrm{d}\boldsymbol{x}_1 \cdots \int \mathrm{d}\boldsymbol{x}_N \\ &\times \langle \boldsymbol{x}_1 \ldots \boldsymbol{x}_N | \exp\left(-\sum_{j=1}^{N} \left(\hat{E}_j^{\mathrm{kin}} / k_{\mathrm{B}} T \right) \right) \exp\left(-\hat{V} / k_{\mathrm{B}} T \right) | \boldsymbol{x}_1 \ldots \boldsymbol{x}_N \rangle \,, \end{aligned} \tag{5.5}$$

where we have used the result

$$e^{\hat{A}+\hat{B}} = e^{\hat{A}} e^{\hat{B}} \tag{5.6}$$

for operators $\hat{A}, \hat{B}$ that commute with each other – which is true, of course, only in the limit $\hbar \to 0$, which is precisely the limit in which quantum mechanics reduces to classical mechanics. If (5.5) were true, one could furthermore use

$$\begin{aligned} \exp\left(-\sum_{j=1}^{N} \hat{E}_j^{\mathrm{kin}} / k_{\mathrm{B}} T \right) &= \prod_{j=1}^{N} \exp\left(-\hat{E}_j^{\mathrm{kin}} / k_{\mathrm{B}} T \right) \\ &= \prod_{j=1}^{N} \exp\left(-\hat{\boldsymbol{p}}_j^2 / 2m k_{\mathrm{B}} T \right) , \end{aligned}$$

and introducing then suitably complete sets in momentum representation $\int \mathrm{d}\boldsymbol{p}_j |\boldsymbol{p}_j\rangle \langle \boldsymbol{p}_j| = \hat{1}$ the kinetic energy terms can simply be evaluated and in the end cancel out from the average in (5.1) if we consider a quantity that depends on the positions of the particles only, since the same expression results from the kinetic energy contributions both in the numerator and in the denominator of (5.1).

As we have emphasized above, details of this reasoning can be found in standard text books on statistical mechanics, and there is no need to dwell on it here. But let us recall what are the physical consequences when we indeed ignore (5.4) and evaluate all averages according to classical rather than quantum statistical mechanics. First of all, we miss spectacular effects which result from the indistinguishability of quantum

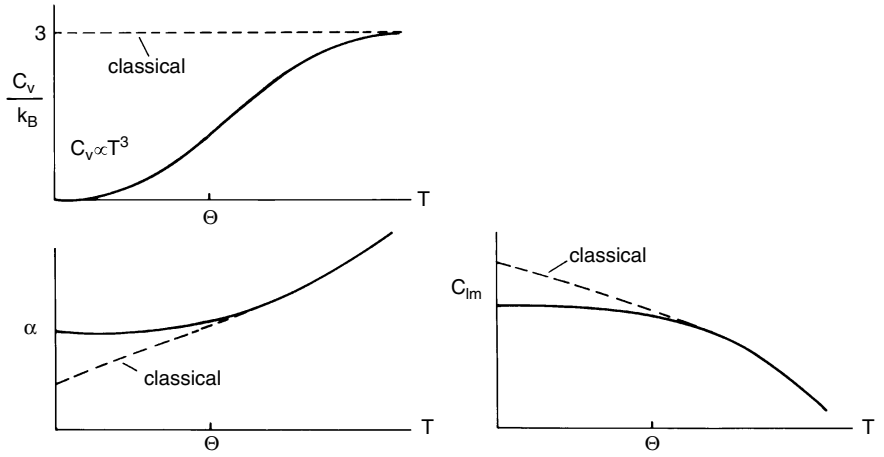


Fig. 5.1 Qualitative sketch of the specific heat C_v (*upper left part*), the lattice constant a (*lower left part*) and an elastic constant $C_{\ell m}$ (*right part*) plotted versus temperature T. The corresponding behavior given by classical statistical mechanics are the *broken straight lines*. Significant differences between classical and quantum statistical mechanics occur for temperatures T below the Debye temperature Θ

particles and the resulting possibility of quantum-mechanical exchange, such as superfluidity and Bose condensation (remember that here we talk about neutral atoms, their electrons being bound to the nuclei – metals with quasi-free electrons and resulting phenomena such as superconductivity and the problem of Anderson localization etc., will not be considered here). But even in cases when we consider relatively heavy atoms (e.g., considering noble gases we exclude helium but wish to deal with neon, argon, xenon), where at low temperatures the fluid–solid transition precludes the occurrence of a suprafluid phase, classical statistical mechanics at low temperatures is severely in error. "Low temperature" means here temperatures comparable to or lower than the Debye temperature Θ, as a treatment of the statistical mechanics of the crystal in terms of phonons shows. Some consequences of quantum mechanics on the low-temperature properties of crystals are sketched qualitatively in Fig. 5.1: while the specific heat per atom of a (harmonic) solid would simply follow the Dulong–Petit law $C_V = 3k_B$ (in $d = 3$ dimensions), it actually vanishes for $T \to 0$ as required by the third law of thermodynamics, and follows the Debye law $C_V \propto T^3$ (in $d = 3$ dimensions) for $T \ll \Theta$; the lattice parameter according to classical statistical mechanics at low T varies linearly with T, which would mean the thermal expansion coefficient would become constant, while in reality the thermal expansion coefficient also vanishes for $T \to 0$, and the same holds for temperature derivatives of elastic constants.

Is there a need at all to study such low-temperature properties with simulations? Isn't it good enough to work out the statistical mechanics based on the lattice dynamical phonon treatment? At this point, it must be emphasized that the simple

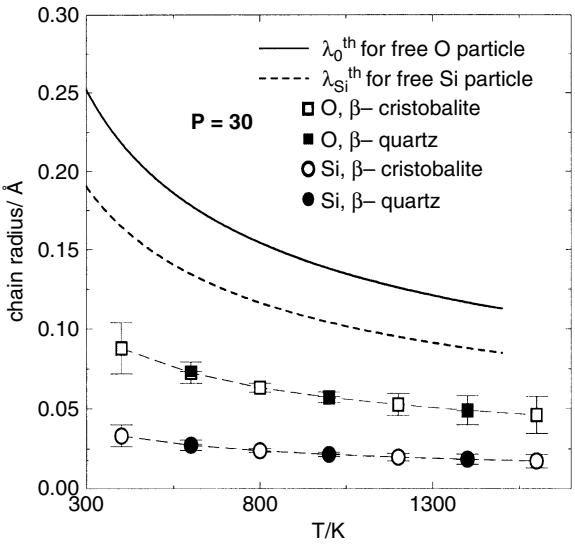


Fig. 5.2 Radii of "ring polymers" representing quantum-mechanically treated silicon and oxygen atoms in two crystalline structures of SiO_2, β-cristobalite and β-quartz, plotted as a function of temperature, using a Trotter number $P = 30$. For comparison, the thermal de Broglie wavelengths for free oxygen and silicon atoms are also shown. (From [5.1])

harmonic approximation for crystals yields the Debye law but it does not yield any temperature dependence of the lattice parameters and the elastic constants at all. In order to account for these temperature dependencies, one has to use at least the self-consistent quasi-harmonic theory. The latter relies on the fact that at nonzero temperature in the NVT ensemble it is not the internal energy minimum which yields the thermal equilibrium but the free energy is a minimum, and at nonzero temperature in a crystal internal energy E and free energy F differ by the entropic contribution ($F = E - TS$) due to the disorder caused by the displacements of the atoms associated with the phonon vibrations. However, we emphasize that this quasi-harmonic theory is not exact due to the neglect of anharmonic terms. While, according to classical statistical mechanics, the harmonic approximation does get exact as $T \to 0$, this is not true if quantum effects are taken into account, due to zero-temperature motions the anharmonicity of the potential always plays some role. These zero-temperature motions are a direct consequence of (5.4) by the resulting Heisenberg uncertainty principle: a particle in a gas would be "spread out" over a linear dimension given by the thermal de Broglie wavelength,

$$\lambda_T = h/\sqrt{2\pi m k_B T}, \tag{5.7}$$

the delocalization of particles around the lattice sites of a crystal (in the potential from their neighbors) may be smaller (see Fig. 5.2 and [5.1]), but it also increases at low temperatures proportional to $T^{-1/2}$, as the thermal de Broglie wavelength does (5.7). On the other hand, Fig. 5.2 demonstrates that for atoms such as Si or O the delocalization of atoms due to zero-point motion at the temperatures of interest is only of the order of 0.1 Å or even smaller, i.e., much less than all interatomic spacings: therefore the neglect of quantum statistics and its consequences (exchange due to

direct overlap of wave functions of particles at neighboring sites) is not a problem in practice.

The precise estimation of lattice parameters, elastic constants and other properties of crystals is of interest in materials science [5.2]. As an example, Fig. 5.3 [5.3] shows the lattice parameter of orthorhombic polyethylene versus temperature, and the fluctuation of the bond angle between three successive carbon atoms. One sees that this lattice parameter does show a behavior as anticipated qualitatively in Fig. 5.1, but even at room temperature the classical and the quantum-mechanical calculations do not yet coincide (note that polyethylene melts at about $T_{\mathrm{m}} = 413\,\mathrm{K}$). Quantum effects are particularly strong in solid polymers, since H and C are such light atoms, and although the bond-angle potential along the C − C chain is rather stiff, there does occur an appreciable zero-point fluctuation in the bond angle. These results [5.3], demonstrating the importance of quantum effects in ordinary polymers such as polyethylene, have come somewhat as a surprise, since polyethylene melts already at $T_{\mathrm{m}} = 413\,\mathrm{K}$, and usually one expects quantum effects to be strong only far below the melting temperature. However, in the case of C_nH_{2n+2} one must take into account the particular anisotropy of the crystal, the covalent forces along the backbone of the chain molecules are very much stronger than the weak van der Waals-type forces between neighboring chain molecules, and the latter forces are only responsible for melting.

Of course, many more problems of the type shown in Figs. 5.1, 5.2 and 5.3 exist in the physics of condensed matter. Particularly interesting are again phase-transition phenomena, and indeed we encounter many transitions in solids from one crystal structure to another driven either by varying the temperature or by varying the pressure. At such structural phase transitions, the local potential experienced by atoms sometimes is of the double-well type: this allows for quantum phenomena which have no classical counterpart at all, such as tunnelling. We also emphasize that interesting quantum effects in condensed-matter physics not only arise from the non-commutativity of the position and momentum operators (5.4), but similar considerations can be made for operators associated with other physical observables as well, e.g., (orbital) angular momentum and spin. Consider the problem of monolayers of adsorbed molecules such as N_2 on graphite [5.4]: In the $\sqrt{3} \times \sqrt{3}$ commensurate superstructure, one may ignore both the translational degrees of freedom and the out-of-plane rotation, and the only degree of freedom that one must consider is the angle φ_i describing the orientation of molecule i in the xy-plane parallel to the substrate. Then the Hamiltonian is (I is the moment of inertia of the molecules, $\hat{L}_{jz}$ is the operator associated the z-component of the angular momentum of molecule j and $\hat{V}$ the intermolecular potential)

$$\hat{\mathcal{H}} = \sum_{j=1}^{N} \frac{\hat{L}_{jz}^2}{2I} + \sum_{i<j} \hat{V}\left(\hat{\varphi}_i, \hat{\varphi}_j\right). \tag{5.8}$$

For this problem the commutation relation analogous to (5.4) reads

$$\left[\hat{L}_{jz}, \hat{\varphi}_k\right] = -i\hbar\delta_{jk}, \tag{5.9}$$

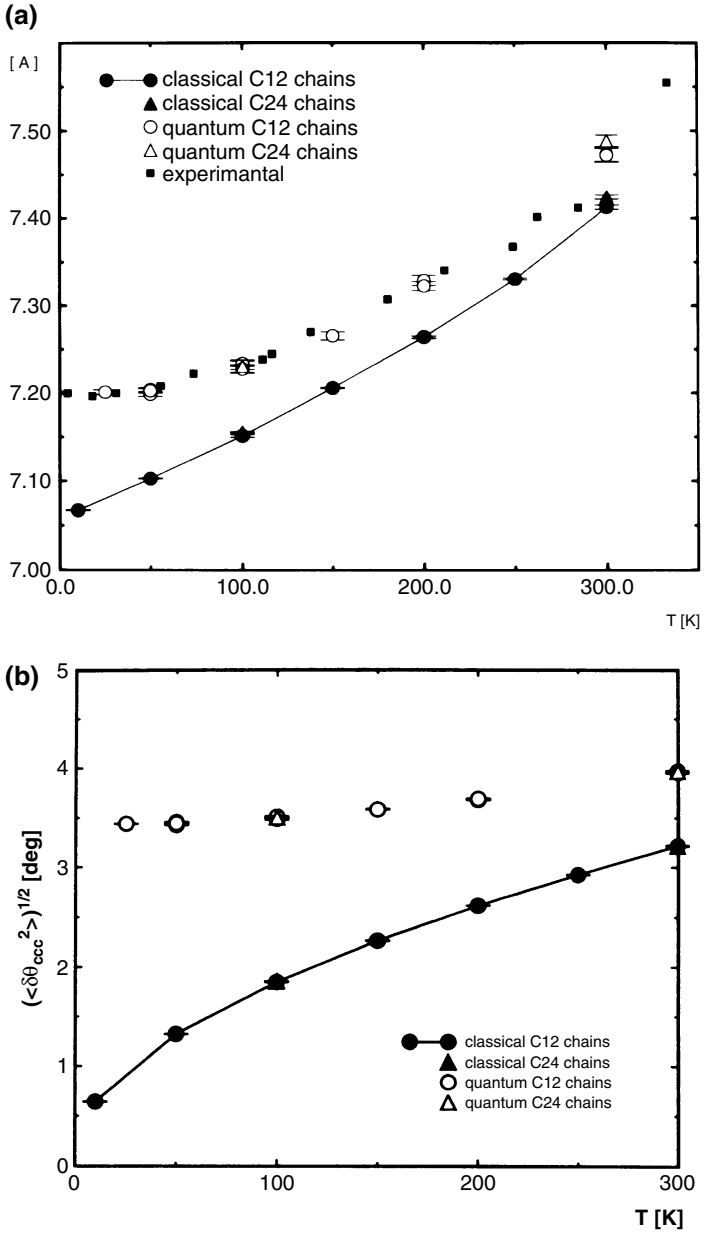


Fig. 5.3 **a** Temperature dependence of the lattice constant a for orthorhombic polyethylene (C_nH_{2n+2}). Results of a path-integral Monte Carlo calculation are compared with the values for a classical system and from experiments. After Martonak et al. [5.3]. **b** Temperature dependence of the average fluctuation $\langle(\delta\Phi_{\mathrm{CCC}})^2\rangle^{1/2}$ of the C − C − C bond angle in polyethylene, according to the classical Monte Carlo simulation (*full dots*), yielding $\langle(\delta\Phi_{\mathrm{CCC}})^2\rangle^{1/2} \propto T^{1/2}$ at low T, and according to path integral Monte Carlo simulations (*open symbols*). Two choices of chain length n are used, $n = 12$ and $n = 24$, respectively. (After [5.3])

and the resulting zero-point vibrations of the angles $\hat{\varphi}_k$ cause an appreciable reduction of the orientational order parameter in comparison with its classical value at low temperature [5.5]. Finally, considering the classical Heisenberg model of Chap. 1 again,

$$\mathcal{H} = -J \sum_{i<j} \hat{\mathbf{S}}_i \cdot \hat{\mathbf{S}}_j - \mu_\mathrm{B} H \sum_{i=1}^{N} \hat{S}_j^z, \tag{5.10}$$

but now rather than unit vectors the $\hat{\mathbf{S}}_i$ are spin operators, satisfying the commutation relation

$$\left[\hat{S}_i^\alpha,\ \hat{S}_j^\beta\right] = i\hbar\varepsilon_{\alpha\beta\gamma}\hat{S}_j^\gamma. \tag{5.11}$$

Since the total magnetization $\hat{M} = \mu_\mathrm{B} \sum_{i=1}^{N} \hat{S}_i^z$ commutes with the Hamiltonian, quantum effects do not lead to a reduction of the magnetization relative to its classical ground state value. But, as spin wave theory shows, the temperature dependence of internal energy and magnetization are very different in the quantum case from the classical case,

$$E(T) = \langle\hat{\mathcal{H}}\rangle = E_0 + \text{const}\, T^{5/2}, \quad \text{quantum case}, \tag{5.12}$$

$$M_0 - M(T) \propto T^{3/2}, \quad \text{quantum case}, \tag{5.13}$$

while

$$E(T) - E_0 \propto T, \quad M_0 - M(T) \propto T, \quad \text{classical case}. \tag{5.14}$$

Again one sees that in the classical case a nonzero specific heat results for $T \to 0$, analogous to the Dulong–Petit result for crystals. And just as the phonon theory of crystals is difficult if one wishes to include anharmonic terms, such terms exist for spin waves (magnons) too and again cannot be accounted for completely by exact analytical methods. Quantum Monte Carlo methods, however, can be applied without such restrictions.

5.2 The Path Integral Quantum Monte Carlo Method

The basic idea of the path integral representation of the partition function [5.6] can already be explained for the simple problem of a single quantum particle in one dimension x in an external potential $\hat{V}(x)$, where (5.3) and (5.2) reduce to

$$\hat{\mathcal{H}} = \hat{E}^\mathrm{kin} + \hat{V} = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + \hat{V}(x), \tag{5.15}$$

$$Z = \int \mathrm{d}x \langle x| \exp(-\hat{\mathcal{H}}/k_\mathrm{B}T)|x\rangle = \int \mathrm{d}x \langle x| \exp(-(\hat{E}_\mathrm{kin} + \hat{V})/k_\mathrm{B}T)|x\rangle. \tag{5.16}$$

The path integral representation of (5.16) can be most easily derived if one recalls the Trotter–Suzuki formula [5.7, 5.8]

$$\exp(\hat{A}+\hat{B}) = \lim_{P\to\infty} [\exp(\hat{A}/P)\exp(\hat{B}/P)]^P, \tag{5.17}$$

which holds for two non-commuting operators $\hat{A}$, $\hat{B}$, which satisfy another operator identity

$$\exp(\hat{A}+\hat{B}) = \exp(\hat{A})\exp(\hat{B})\exp\left(-\frac{1}{2}[\hat{A},\hat{B}]\right), \tag{5.18}$$

when the commutator of the operators $\hat{A}$, $\hat{B}$ is a complex number c, i.e., $[\hat{A},\hat{B}] = c$. Equation (5.18) can be easily derived from systematic Taylor expansions of the exponential function and should be familiar to the reader from elementary text books on quantum mechanics. If we now apply (5.18) to the operator $\exp(\hat{A}'+\hat{B}')$ where $\hat{A}' = \hat{A}/P$, $\hat{B}' = \hat{B}/P$, we recognize that the term $[\hat{A}',\hat{B}']$ that appears on the right-hand side in the last exponential is of order P^{-2},

$$[\hat{A}',\hat{B}'] = c/P^2, \tag{5.19}$$

and thus it is plausible that in the limit $P\to\infty$ this correction can be neglected, and hence (5.17) results. In our case of the particle in the external potential we hence use

$$\begin{aligned}&\exp\left[-(\hat{E}^{\mathrm{kin}}+\hat{V})/k_{\mathrm{B}}T\right]\\ &= \lim_{P\to\infty}\left[\exp(-\hat{E}^{\mathrm{kin}}/k_{\mathrm{B}}TP)\exp\left(-\hat{V}/k_{\mathrm{B}}TP\right)\right]^P.\end{aligned} \tag{5.20}$$

Using (5.20), the partition function (5.16) becomes

$$\begin{aligned}Z = \lim_{P\to\infty}&\int \mathrm{d}x_1\int \mathrm{d}x_2\cdots\\ &\times\int \mathrm{d}x_P\langle x_1|\exp(-E^{\mathrm{kin}}/k_{\mathrm{B}}TP)\exp(-\hat{V}/k_{\mathrm{B}}TP)|x_2\rangle\\ &\times\langle x_2|\exp(-\hat{E}^{\mathrm{kin}}/k_{\mathrm{B}}TP)\exp(-\hat{V}/k_{\mathrm{B}}TP)|x_3\rangle\langle x_3|\cdots|x_P\rangle\\ &\times\langle x_P|\exp(-\hat{E}^{\mathrm{kin}}/k_{\mathrm{B}}TP)\exp(-\hat{V}/k_{\mathrm{B}}TP)|x_1\rangle.\end{aligned} \tag{5.21}$$

The matrix elements appearing in (5.21) can be worked out, this is just an exercise in elementary quantum mechanics, and the result is

$$\begin{aligned}&\langle x|\exp(-\hat{E}^{\mathrm{kin}}/k_{\mathrm{B}}TP)\exp(-\hat{V}/k_{\mathrm{B}}TP)|x'\rangle\\ &= \left(\frac{mk_{\mathrm{B}}TP}{2\pi\hbar^2}\right)^{1/2}\exp\left[-\frac{mk_{\mathrm{B}}TP}{2\hbar^2}(x-x')^2\right]\exp\left[-\frac{V(x)+V(x')}{2k_{\mathrm{B}}TP}\right].\end{aligned} \tag{5.22}$$

Thus the partition function becomes

$$
\begin{aligned}
Z &= \lim_{P\to\infty} Z_P, \\
Z_P &= \left(\frac{mk_{\mathrm{B}}TP}{2\pi\hbar^2}\right)^{P/2} \int \mathrm{d}x_1 \int \mathrm{d}x_2 \cdots \\
&\quad \times \int \mathrm{d}x_p \exp\left\{-\frac{1}{k_{\mathrm{B}}T}\left[\frac{\kappa}{2}\sum_{s=1}^{P}(x_s - x_{s+1})^2 + \frac{1}{P}\sum_{s=1}^{P} V(x_s)\right]\right\},
\end{aligned}
\tag{5.23}
$$

where the boundary condition $x_{P+1} = x_1$ is implied, and as an abbreviation an effective spring constant κ was introduced,

$$
\kappa = mP(k_{\mathrm{B}}T/\hbar)^2. \tag{5.24}
$$

From (5.23) one immediately recognizes that Z_P can be considered as the partition function of a problem in classical statistical mechanics with P degrees of freedom, namely a harmonic chain in an external potential $V(x_s)/P$. In this way we have a problem of quantum statistical mechanics (5.16) to which standard Monte Carlo methods, as described in the first chapter of this book, are readily applied. Of course, in practice we will work with several choices of large but finite P, in order to carry out the extrapolation $P \to \infty$ numerically.

This approach can be generalized straightforwardly to N particles interacting with each other according to the potential $\hat{V}(\hat{\boldsymbol{r}}_i - \hat{\boldsymbol{r}}_j)$ in three dimensions (5.1)–(5.3) if we disregard the indistinguishability and statistics of the particles (later on this restriction will be removed of course). Using steps analogous to those which lead from (5.21) to (5.23), use of (5.20) now yields for (5.3) the result

$$
\begin{aligned}
Z_P &= \left(\frac{mk_{\mathrm{B}}TP}{2\pi\hbar^2}\right)^{3NP/2} \\
&\quad \times \int \mathrm{d}\boldsymbol{r}_1^{(1)} \cdots \int \mathrm{d}\boldsymbol{r}_1^{(P)} \int \mathrm{d}\boldsymbol{r}_2^{(1)} \cdots \mathrm{d}\boldsymbol{r}_2^{(P)} \cdots \int \mathrm{d}\boldsymbol{r}_N^{(1)} \cdots \int \mathrm{d}\boldsymbol{r}_N^{(P)} \\
&\quad \times \exp\left\{-\frac{1}{k_{\mathrm{B}}T}\left[\frac{\kappa}{2}\sum_{i=1}^{N}\sum_{s=1}^{P}\left(\boldsymbol{r}_i^{(s)} - \boldsymbol{r}_i^{(s+1)}\right)^2\right.\right. \\
&\qquad \left.\left. + \frac{1}{P}\sum_{i<j}\sum_{s=1}^{P} V\left(|\boldsymbol{r}_i^{(s)} - \boldsymbol{r}_j^{(s)}|\right)\right]\right\}.
\end{aligned}
\tag{5.25}
$$

Equation (5.25) can be interpreted as a melt of cyclic chains ("ring polymers") with harmonic springs connecting neighboring beads along the chains, but the interactions among the beads are rather uncommon: while in a physical melt of ring polymers in principle every bead can interact with every other bead in the system, here only beads with the same Trotter index s are interacting (the coordinate along

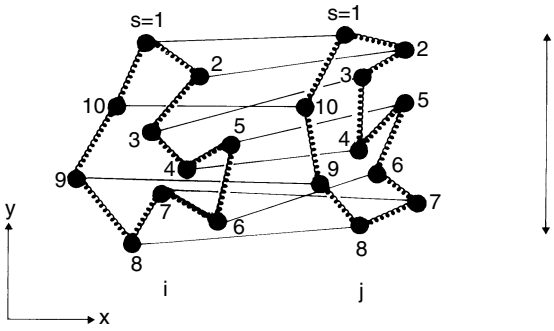


Fig. 5.4 Schematic representation of two interacting quantum particles i, j in two dimensions: each particle (i) is represented by a "ring polymer" composed of $P = 10$ effective monomers at positions $\boldsymbol{r}_i^{(s)}$, with $s = 1, 2, \ldots, P$. Harmonic springs (of strength κ) only connect "monomers" of the same "polymer" while interatomic forces join different monomers with the same Trotter index s, Indicated by the *thin straight lines*. In the absence of such interactions, the size of such a ring polymer would be of the order of the thermal de Broglie wavelength $\lambda_T = h/\sqrt{2\pi m k_\mathrm{B} T}$, where h is Planck's constant

the Trotter index s plays the role of imaginary time $\tau = s\Delta\tau$ [5.6] and hence there is only an interaction between particles that belong to the same "timeslice" ($\Delta\tau$). This situation is sketched in Fig. 5.4 for two interacting particles.

By (5.25), the problem of evaluating the partition function, (5.2), within the framework of quantum statistical mechanics has indeed been reduced to an equivalent problem in classical statistical mechanics, although one must be careful since the effective Hamiltonian $\mathcal{H}p$ of the $N \times P$ particles in Z_P

$$Z_P = \left(\frac{mk_\mathrm{B}TP}{2\pi\hbar^2}\right)^{3NP/2} \int \mathrm{d}\boldsymbol{r}_1^{(1)} \cdots \int \mathrm{d}\boldsymbol{r}_N^{(P)} \exp\left\{-\frac{1}{k_\mathrm{B}T}\mathcal{H}_P\right\} \tag{5.26}$$

depends explicitly on temperature (via the spring constant κ, see (5.24)) [5.6, 5.9–5.12]. Thus Monte Carlo simulations can be applied rather straightforwardly to estimate corresponding averages [5.9–5.12]

$$\langle A\rangle_P = Z_P^{-1} \int \mathrm{d}\boldsymbol{r}_1^{(1)} \cdots \int \mathrm{d}\boldsymbol{r}_N^{(P)} \exp\left\{-\frac{1}{k_\mathrm{B}T}\mathcal{H}_P\right\} A. \tag{5.27}$$

Let us now briefly discuss the physical interpretation of these results. If the potential in (5.25) could be neglected completely, we could infer from the equipartition theorem of classical statistical mechanics that the energy carried by each spring is (in d dimensions)

$$\frac{\kappa}{2}\left\langle\left(\boldsymbol{r}_i^{(s)} - \boldsymbol{r}_i^{(s+1)}\right)^2\right\rangle = \frac{d}{2}k_\mathrm{B}T. \tag{5.28}$$

Using (5.24) and (5.28), we conclude that the average mean square distance between neighboring particles in a ring polymer is

$$\ell^2 \equiv \left\langle \left(\boldsymbol{r}_i^{(s)} - \boldsymbol{r}_i^{(s+1)} \right)^2 \right\rangle = dk_{\mathrm{B}}T/\kappa = \hbar^2 d/(mk_{\mathrm{B}}TP). \tag{5.29}$$

Of course, this result does not depend on s, all particles in a ring polymer are equivalent. Now the gyration radius of a ring polymer containing P monomers for large P scales as

$$\left\langle R_g^2 \right\rangle = \ell^2 P/12 = (d/12)(\hbar^2/(mk_{\mathrm{B}}T)). \tag{5.30}$$

As could have been expected, this is of the same order as the square of the thermal de Broglie wavelength, $\lambda_T^2 = h^2/(2\pi mk_{\mathrm{B}}T)$, and the "Trotter number" P has cancelled out from this result. As is well known from the quantum statistical mechanics of an ideal, noninteracting gas, integrating out the momenta the Heisenberg uncertainty principle $(\Delta p)^2(\Delta x)^2 \approx h^2$ requires that the delocalization of a particle in space is $\langle(\Delta x)^2\rangle \approx h^2/\langle(\Delta p)^2\rangle \approx \hbar^2/(mk_{\mathrm{B}}T)$, omitting prefactors of order unity here. Equation (5.30) thus again illustrates that quantum effects become more important the lower the temperature and the lighter the particle (i.e., the smaller the mass m). Of course, (5.29) no longer is accurate when interactions among the particles occur, but qualitatively it still predicts correctly the order of magnitude of the quantum mechanical delocalization due to zero-point motions at low temperatures; Fig. 5.2 gives an explicit example for the case of SiO_2 [5.1].

Now it is clear that in order to take into account quantum effects correctly one has to perform the extrapolation towards $P \to \infty$ (see (5.23)). From (5.17) to (5.20) one can infer that the quantum corrections of physical observables scale as P^{-2} (see (5.19)). This result implies that one should compute observables for several values of P and try an extrapolation of the results as a function of P^{-2}. Of course, in general it is a nontrivial question to judge how large P must be chosen in order to reach the asymptotic scaling limit. From (5.5) it is clear that the distance between effective monomers in the ring polymer scales as $\ell \propto \hbar/\sqrt{mk_{\mathrm{B}}TP}$. We need to keep this distance fixed at a value that is small in comparison to the length scale on which the potential acting on the particles changes appreciably. In order to keep ℓ fixed it thus obviously is necessary to keep the product $mk_{\mathrm{B}}TP$ fixed – the lower the temperature, the larger P must be chosen. In practice, there does not exist a simple recipe that tells us how large P has to be in a specific case; rather one has to find the appropriate range of P values by trial and error. Figure 5.5 shows that in favorable cases rather small values of P suffice to reach the asymptotic limit [5.13] where the scaling of data linearly with P^{-2} actually is observed. This figure also demonstrates that PIMC is able to identify typical quantum-mechanical effects such as "isotope effects": the two isotopes ^{20}Ne and ^{22}Ne of the Lennard–Jones system differ only by their mass – in classical statistical mechanics there would be no difference in static properties whatsoever. However, Fig. 5.5 shows there is a clear distinction between the lattice constants of the two isotopes, and the difference observed in the simulation in fact is rather close to the value found in experiments [5.14]. However, other examples exist

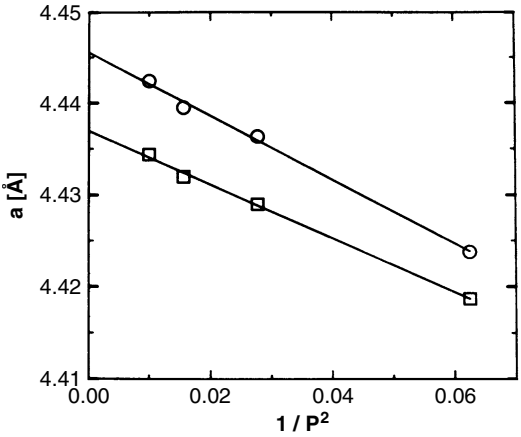


Fig. 5.5 Trotter scaling plot for the lattice parameter a of solid neon. The *upper curve* corresponds to ^{20}Ne at $T = 16\,\mathrm{K}$, the *lower curve* to ^{22}Ne at $T = 16\,\mathrm{K}$. (From [5.13])

when even Trotter numbers as large as $P = 100$ are not large enough to ensure that the asymptotic limit of the P^{-2} scaling has been reached.

As mentioned above, the treatment so far has completely neglected the effects of quantum statistics. This approximation is ok for crystals, ensuring furthermore that the typical inter-particle distance is large in comparison with the linear dimension of the ring polymers. As expected, this approximation breaks down when the linear dimension of the ring polymer describing a particle $\left(\sqrt{\langle R_g^2 \rangle}\right)$ becomes comparable to inter-particle distances: then the wave functions show appreciable overlap and the effects of quantum statistics need to be properly taken into account. This problem needs to be handled when one wishes to treat quantum crystals such as solid ^{3}He or solid ^{4}He, as well as the corresponding quantum fluids [5.11, 5.12]. Here we shall only discuss Bose systems (such as ^{4}He). Only totally symmetric eigenfunctions contribute to the density matrix, and hence we introduce a permutation operator $\hat{P}$ such that $\hat{P}\boldsymbol{R}$ is a permutation of particle labels if we use the shortened notation $\boldsymbol{R} \equiv (\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_N)$ for the set of particle coordinates. Then we have for any eigenfunction $\Phi_\alpha(\hat{R})$

$$\hat{P}\Phi_\alpha(\boldsymbol{R}) = \frac{1}{N!}\sum_P \Phi(\hat{P}\boldsymbol{R}), \tag{5.31}$$

and the partition function for a Bose system therefore takes the form

$$Z_{\mathrm{Bose}} = \left(\frac{mk_\mathrm{B}TP}{2\pi\hbar^2}\right)^{dNP/2} \frac{1}{N!}\int \mathrm{d}\boldsymbol{r}_1^{(1)} \cdots \int \mathrm{d}\boldsymbol{r}_N^{(P)} \exp(-\mathcal{H}_P/k_\mathrm{B}T). \tag{5.32}$$

Now the boundary condition is not $\boldsymbol{r}_i^{(P+1)} = \hat{r}_i^{(1)}$ as it was in (5.25) and (5.26), but one has to obey only the condition that

$$\hat{P}\boldsymbol{R}^{(P+1)} = \boldsymbol{R}^{(1)}. \tag{5.33}$$

This means that paths are allowed to close on any permutation of their starting positions, and contributions from all $N!$ closures are summed over in the partition function. At high temperatures the contribution from the identity permutation will dominate, while at zero temperature all permutations contribute equally. In the classical isomorphic polymer system, this means that "cross-links" can form between chains and open up again; of course, this has nothing to do with the chemical kinetics of cross-linking and polymerization in real polymers. Thus, a two-atom system with P effective monomers per ring polymer can be in two possible permutation states: either two separate ring polymers, each with P springs (this is the situation described in Fig. 5.4) or one larger ring polymer with $2P$ springs.

It is illuminating to ask what superfluidity means in this formalism [5.15], since this actually occurs in ^{4}He: A macroscopic polymer is formed which involves of the order of N atoms and stretches over the entire system. From Fig. 5.4, it is clear that this "cross-linking" among ring polymers can set in only when the linear dimension of a ring polymer becomes of the same order as the "ring polymer spacing"; from this argument one can get a rough estimate of the superfluid transition temperature, by putting the thermal de Broglie wavelength $\lambda_T = h/\sqrt{2\pi m k_\mathrm{B} T}$ equal to the classical inter-particle spacing, $\varrho^{-1/d}$ in d dimensions, ϱ being the density. The "degeneracy temperature" T_D found from $\lambda_T = \varrho^{-1/d}$ is $T_\mathrm{D} = \varrho^{2/d} h^2/(2\pi k_\mathrm{B} m)$, and this sets the temperature scale on which important quantum effects occur.

In practice, use of (5.25) and (5.30) would not work for the study of superfluidity in ^{4}He: although the formalism in principle is exact, too large values of P would be required in order to obtain reasonable results. In order to make progress one must not use the so-called "primitive action" defined in (5.25) but must use the so-called "improved actions" for $\mathcal{H}_P$. We refer the reader to the original literature for details [5.11].

5.3 Quantum Monte Carlo for Lattice Models

One follows again the strategy to decompose the Hamiltonian of the considered model $\mathcal{H}$ into two parts, $\hat{\mathcal{H}} = \hat{\mathcal{H}}_1 + \hat{\mathcal{H}}_2$, that can be diagonalized separately so that the use of the Trotter–Suzuki formula [5.7, 5.8] is helpful, see (5.17),

$$\exp[-(\hat{\mathcal{H}}_1 + \hat{\mathcal{H}}_2)/k_\mathrm{B}T] = \lim_{P\to\infty} [\exp(-\hat{\mathcal{H}}_1/k_\mathrm{B}TP)\exp(-\hat{\mathcal{H}}_2/k_\mathrm{B}TP)]^P . \quad (5.34)$$

Unfortunately, there is no general recipe how this splitting of the Hamiltonian into parts is best done in practice – what is best depends very much on the model that is considered. Therefore many different variants of this approach can be found in the literature [5.16–5.18]; we hence cannot attempt to explain all these various methods, but only attempt to convey the spirit of the general approach here. At this point, we also mention that it is possible to consider higher-order decompositions of

$\hat{\mathcal{H}}$, where application of the Trotter formula with a finite Trotter number P does not imply a scaling of the error as P^{-2} but according to a higher power of $1/P$ [5.19].

As a first example, we treat the one-dimensional Ising model in a transverse field of strength $H_\perp$, taking [5.20]

$$\hat{\mathcal{H}}_1 = -J \sum_{i=1}^{N} \hat{\sigma}_i^z \hat{\sigma}_{i+1}^z, \quad \hat{\mathcal{H}}_2 = -H_\perp \sum_{i=1}^{N} \hat{\sigma}_i^x, \tag{5.35}$$

where $\hat{\sigma}_i^\alpha (\alpha = x, y, z)$ denote the Pauli spin matrices at lattice site i. Periodic boundary conditions $\hat{\sigma}_{N+1}^\alpha = \hat{\sigma}_1^\alpha$ are assumed as usual. As a state representation let us use eigenstates of $\hat{\sigma}_z$ and label them by Ising spin variables, $s = \pm 1$, i.e.,

$$\hat{\sigma}^z |s\rangle = s|s\rangle. \tag{5.36}$$

Of course, $\hat{\mathcal{H}}_1$ is diagonal in this representation, while $\hat{\mathcal{H}}_2$ is not. Now the Pth approximant Z_P to the partition function can be written as

$$\begin{aligned} Z_P &= \mathrm{Tr} \left\{ \exp(-\hat{\mathcal{H}}_1/k_\mathrm{B}TP) \exp(-\hat{\mathcal{H}}_2/k_\mathrm{B}TP) \right\}^P \\ &= \sum_{\left\{S_i^{(s)}\right\}} \prod_{s=1}^{P} \prod_{i=1}^{N} \exp\left[\frac{J}{k_\mathrm{B}TP} S_i^{(s)} S_{i+1}^{(s)} \right] \left\langle S_i^{(s)} \right| \exp\left(\frac{H_\perp \hat{\sigma}_i^x}{k_\mathrm{B}TP} \right) \left| S_i^{(s+1)} \right\rangle. \end{aligned} \tag{5.37}$$

In this trace we have to take periodic boundary conditions in the imaginary time direction as well, $S_i^{(P+1)} = S_i^{(1)}$. Using the results for the Pauli spin operators, one easily obtains the following result for the matrix element in (5.37)

$$\langle s| \exp(a\hat{\sigma}^x)|s'\rangle = \left(\frac{1}{2} \sinh 2a \right)^{1/2} \exp\left(\frac{1}{2} \log \coth a \right) ss', \tag{5.38}$$

where a is an abbreviation for $H_\perp/(k_\mathrm{B}TP)$. Using (5.38) in Z_P one obtains an expression that is formally equivalent to the partition function of an anisotropic two-dimensional Ising model in the absence of any fields, namely

$$\begin{aligned} Z_P = &\left[\frac{1}{2} \sinh(2H_\perp/k_\mathrm{B}TP) \right]^{PN/2} \\ &\times \sum_{\left\{S_i^{(k)}\right\}} \exp\left[\sum_{k=1}^{P} \sum_{i=1}^{N} \left(K_P S_i^{(k)} S_i^{(k+1)} + \frac{J}{k_\mathrm{B}TP} S_i^{(k)} S_{i+1}^{(k)} \right) \right], \end{aligned} \tag{5.39}$$

with a coupling constant K_P in the "Trotter direction" that depends both on the temperature T and the linear dimension P in this direction,

$$K_P = \frac{1}{2} \log\{\coth(H_\perp / k_B T P)\}. \tag{5.40}$$

Equation (5.40) is analogous to the coupling with the spring constant κ in (5.24). Again it turns out necessary to choose P such that one works in the limit of large K_P, i.e., one must have $k_B T P / H_\perp \gg 1$, the lower the temperature the larger the Trotter dimension P must be. As in the off-lattice case, the original interaction (here the exchange interaction J), acts only between spins with the same Trotter index (Denoted as k here).

The partition function of the two-dimensional Ising square lattice can be solved exactly also for anisotropic exchange couplings, and hence there is no need to deal with this problem by Monte Carlo methods. However, the same method as shown in (5.35)–(5.40) straightforwardly applies to higher-dimensional Ising models with transverse fields as well – always the quantum effects lead to the occurrence of this extra dimension, and the linear dimension P in this direction needs to be extrapolated to infinity in order to render this method as an exact one. In practice, the recipe is to carry out a series of simulations for finite values of P, and extrapolate physical properties as functions of P^{-2} towards $P \to \infty$.

As a second and physically more interesting example, where the Trotter formalism is applied to a spin problem on a lattice, we have the anisotropic Heisenberg chain with spin quantum number $s = 1/2$. The Hamiltonian of this model is given by (periodic boundary conditions again being implied)

$$\mathcal{H} = \sum_{i=1}^{N} (J_x \hat{S}_i^x \hat{S}_{i+1}^x + J_y \hat{S}_i^y \hat{S}_{i+1}^y + J_z \hat{S}_i^z \hat{S}_{i+1}^z). \tag{5.41}$$

There have been several distinct ways in which the quantum Hamiltonian can be split into parts such that the Trotter formula (5.17) can be applied in a useful way. We describe here only the procedure first suggested by Suzuki and by Barma and Shastry [5.21, 5.22],

$$\hat{\mathcal{H}} = \hat{\mathcal{H}}_0 + \hat{\mathcal{H}}_A + \hat{\mathcal{H}}_B, \tag{5.42}$$

where

$$\hat{\mathcal{H}}_0 = -\sum_{i=1}^{N} J_z \hat{S}_i^z \hat{S}_{i+1}^z, \quad \hat{\mathcal{H}}_A = \sum_{i \text{ odd}} \hat{\mathcal{H}}_i, \quad \hat{\mathcal{H}}_B = \sum_{i \text{ even}} \hat{\mathcal{H}}_i, \tag{5.43}$$

where $\hat{\mathcal{H}}_i$ is the local transverse part of the Hamiltonian,

$$\hat{\mathcal{H}}_i = -\left(J_x \hat{S}_i^x \hat{S}_{i+1}^x + J_y \hat{S}_i^y \hat{S}_{i+1}^y \right). \tag{5.44}$$

We apply the Trotter formula to obtain the Pth approximant Z_P of the partition function in the following form,

$$Z_P = \mathrm{Tr}\left(e^{-\hat{\mathcal{H}}_0/2k_\mathrm{B}TP} e^{-\hat{\mathcal{H}}_A/k_\mathrm{B}TP} e^{-\hat{\mathcal{H}}_0/2k_\mathrm{B}TP} e^{-\hat{\mathcal{H}}_B/k_\mathrm{B}TP}\right)^P . \tag{5.45}$$

As in the previous case, we use eigenstates of $\hat{S}^z$ and hence of the Hamiltonian $\hat{\mathcal{H}}_0$, the Ising-like part. We now insert into the above trace operation altogether $2P$ complete sets of such states in such a way that there is just one complete set between each term $e^{-\hat{\mathcal{H}}_A/k_\mathrm{B}TP}$, $e^{-\hat{\mathcal{H}}_B/k_\mathrm{B}TP}$.

$$Z_P = \underset{\{S_i^{(k)}\}}{\mathrm{Tr}} \exp\left\{-\frac{1}{2k_\mathrm{B}TP}\sum_{k=1}^{2P}\mathcal{H}_0^{(k)} - \frac{1}{k_\mathrm{B}T}\sum_{i\in A}\sum_{k=1}^{2P} h(i,k) - \frac{1}{k_\mathrm{B}T}\sum_{i\in B}\sum_{k=1}^{2P} h(i,k)\right\}, \tag{5.46}$$

with

$$\exp\left[-h(i,k)/k_\mathrm{B}T\right] = \left\langle S_i^{(k)} S_{i+1}^{(k)}\right| \exp(-\mathcal{H}_i/k_\mathrm{B}TP)\left|S_i^{(k,i)} S_{i+1}^{(k)}\right\rangle. \tag{5.47}$$

Also the spins $S_i^{(k)}$ have values $S_i^{(k)} = \pm 1/2$. Equation (5.46) can be interpreted as the partition function of a lattice of size $N \times 2P$ with periodic boundary conditions in both directions and a very anisotropic interaction: these are just the two-spin couplings described by $\mathcal{H}_0^{(k)} = -\sum_{i=1}^{N} J_z S_i^{(k)} S_{i+1}^{(k)}$ in the real space direction, and temperature-dependent four-spin couplings on alternating elementary plaquettes, which couple neighboring sites in both real space and the Trotter direction. This one can recognize from (5.47), which defines the four-spin couplings implicitly.

For more details on this problem defined by (5.41) and results obtained in numerical studies of (5.46) by Monte Carlo methods we refer to the literature [5.20, 5.23]. Here we turn to a very brief discussion of other models that have been intensively studied, that involve fermionic degrees of freedom explicitly. The simplest case is spinless fermions in one dimension [5.20, 5.23], with the Hamiltonian

$$\hat{\mathcal{H}} = -t\sum_{i=1}^{N}(\hat{c}_i^+\hat{c}_{i+1} + \hat{c}_{i+1}^+\hat{c}_i) + V\sum_{i=1}^{N}\hat{n}_i\hat{n}_{i+1}. \tag{5.48}$$

As is well known, the fermion operators $\hat{c}_i^+$, $\hat{c}_i$ create (or annihilate) a particle at lattice site i, and satisfy the anticommutation relation

$$[\hat{c}_i^+, \hat{c}_j]_+ = \hat{c}_i^+\hat{c}_j + \hat{c}_j\hat{c}_i^+ = \delta_{ij}. \tag{5.49}$$

The particle number operator

$$\hat{n}_i \equiv \hat{c}_i^+\hat{c}_i \tag{5.50}$$

has only the two eigenvalues $n_i = 0$ or $n_i = 1$,

$$\hat{n}_i|1\rangle = 1|1\rangle, \quad \hat{n}_i|0\rangle = 0|0\rangle = o, \tag{5.51}$$

expressing the Pauli principle that never can two fermions be in the same state. The total number of particles in the system is then related to the operator

$$\hat{\mathcal{N}} = \sum_{i=1}^{N} \hat{n}_i, \tag{5.52}$$

and the particle density is defined as

$$\hat{\varrho} = \hat{\mathcal{N}}/N. \tag{5.53}$$

In (5.48) one can choose the hopping energy t as unity, leaving V as the only nontrivial energy scale in the model. Since $\hat{c}_i^+|n_i\rangle = |1\rangle$ if $n_i = 0$, $\hat{c}_{i+1}|n_{i+1}\rangle = |0\rangle$ if $n_{i+1} = 1$, the term $\hat{c}_i^+ c_{i+1}$ in (5.48) yields a non-vanishing contribution if a particle is destroyed at site $i+1$ and simultaneously a particle is created at site i, which physically may be interpreted as a hopping process of the particle from site $i+1$ to site i.

It turns out that the present model, (5.48), can be essentially mapped to the previous model, (5.41), by a clever transformation: this approach is one of the standard tricks to deal with quantum problems, by which one tries to make the problem more tractable! Thus, one first defines spin-raising ($\hat{\sigma}_\ell^+$) and spin-lowering ($\hat{\sigma}_\ell^-$) operators in terms of the Pauli matrices $\hat{\sigma}_\ell^\alpha$ ($\alpha = x, y, z$),

$$\hat{\sigma}_\ell^+ = (\hat{\sigma}_\ell^x + i\hat{\sigma}_\ell^y)/2, \quad \hat{\sigma}_\ell^- = (\hat{\sigma}_\ell^x - i\hat{\sigma}_\ell^y)/2. \tag{5.54}$$

Now the fermion operators can be expressed in terms of the operators $\hat{\sigma}_\ell^+$, $\hat{\sigma}_\ell^-$, $\hat{\sigma}_\ell^z$ by the so-called Jordan–Wigner transformation,

$$\hat{c}_\ell^+ = \hat{\sigma}_\ell^+ \exp\left[\frac{i\pi}{2}\sum_{p=1}^{\ell-1}\left(1+\hat{\sigma}_p^z\right)\right], \quad \hat{c}_\ell^- = \hat{\sigma}_\ell^- \exp\left[-\frac{i\pi}{2}\sum_{p=1}^{\ell-1}\left(1+\hat{\sigma}_p^z\right)\right]. \tag{5.55}$$

While this nonlocal transformation looks fairly complicated, the resulting Hamiltonian becomes rather simple, if we neglect boundary terms, which are unimportant for $N \to \infty$,

$$\hat{\mathcal{H}} = -\frac{t}{2}\sum_{i=1}^{N}\left(\hat{\sigma}_i^x\sigma_{i+1}^x + \hat{\sigma}_i^y\hat{\sigma}_{i+1}^y\right) - \frac{V}{2}\sum_{i=1}^{N}\left(\hat{\sigma}_i^z\hat{\sigma}_{i+1}^z + 2\hat{\sigma}_i^z + 1\right). \tag{5.56}$$

This problem can be solved by the method described for (5.41) in (5.42)–(5.51), or a similar decomposition [5.20, 5.24]. Here we do not discuss further the methodological aspects of this problem, but rather show results [5.24] for the structure factor

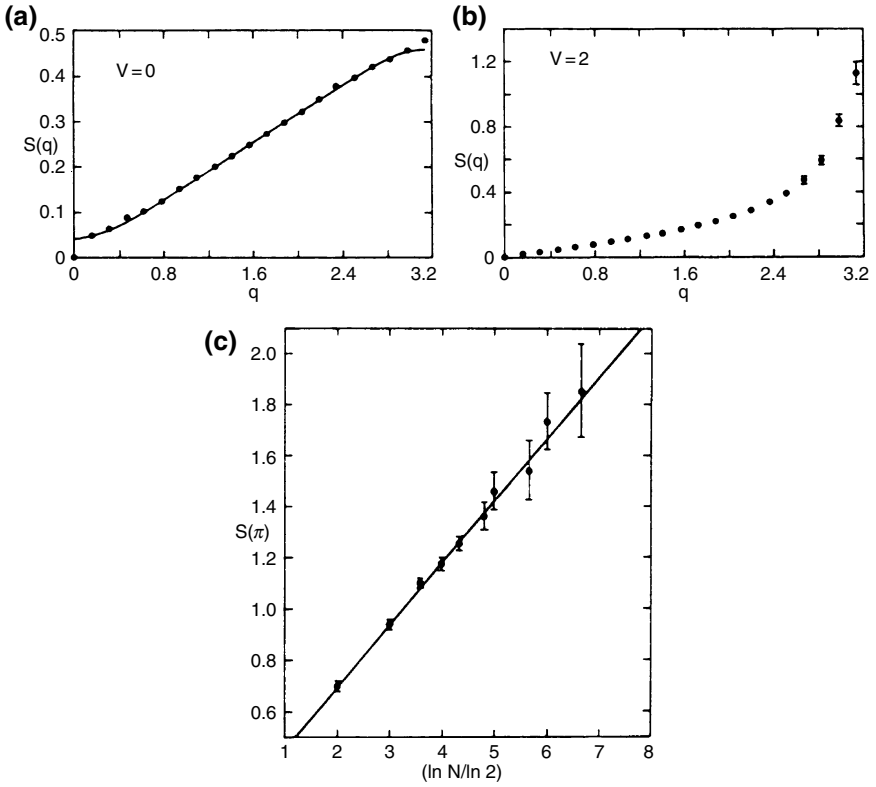


Fig. 5.6 **a** *Points* showing Monte Carlo data for the structure factor $S(q)$ versus q, for the case of $N = 40$ and 20 non-interacting electrons ($t = 1$, $V = 0$) at a low temperature, $k_B T = 1/4$. *Solid line* is the analytical solution, which can trivially be found for this non-interacting system. **b** Monte Carlo results for the structure factor $S(q)$ versus q, for the case of $N = 40$ but 20 interacting electrons ($t = 1, V = 2, k_B T = 1/4$). Note the difference in scale between parts **a** and **b**. **c** Maximum structure factor for the half-filled case ($\langle \hat{\varrho} \rangle = 1/2$), $S(q_\pi)$, plotted versus the logarithm of the lattice size for $t = 1$, $V = 2$, $k_B T = 1/4$. (From [5.24])

$$S_T(q) = \sum_{j=1}^{N} \left(\langle \hat{n}_i \hat{n}_{i+j} \rangle - \langle \hat{n}_i \rangle \langle \hat{n}_{i+j} \rangle \right) \cos(jqa), \tag{5.57}$$

where a is the lattice spacing and q the wave number (Fig. 5.6). One can see that for such fermion models even in $d = 1$ dimensions nontrivial results are obtained. But even in this case statistical errors at low temperatures become appreciable already for moderately large N (the largest N included in Fig. 5.6 is $N = 100$), and the finite-size behavior needs to be carefully analyzed (note the logarithmic scaling with N).

One of the most famous models for interacting electrons on a lattice is the single-band Hubbard model [5.25]

$$\hat{\mathcal{H}} = -t \sum_{\langle i,j \rangle} (\hat{c}^+_{j,\sigma} \hat{c}_{i,\sigma} + \hat{c}^+_{i,\sigma} \hat{c}_{j,\sigma}) + U \sum_i \hat{n}_{i\downarrow} \hat{n}_{i\uparrow}, \tag{5.58}$$

where now $\hat{c}^+_{i,\sigma}(\hat{c}_{i,\sigma})$ denotes a creation (annihilation) operator for a fermion spin σ at lattice site i, with $\sigma = \uparrow$ or $\sigma = \downarrow$ denoting the two orientations of the electron spin. Comparing (5.58) with (5.48), we note that we have made two generalizations: from one dimension to general dimension, but the hopping is still restricted to nearest neighbor pairs $\langle i, j \rangle$; and the fermions are treated as spin $\frac{1}{2}$ particles, as they should be, rather than ignoring the spin, as done in (5.48). Electrons are negatively charged, of course, and so in principle we should have a Coulomb interaction: as a crude approximation, all interactions are neglected apart from the on-site interaction between two electrons with opposite spin on the same site in (5.48); for the case of "spinless fermions", at most one particle per site is possible, and hence the simplest choice of interaction is a nearest-neighbor interaction. Although the model (5.58) is not a realistic description of any physical system, it still captures some essential features of the physics of strongly correlated electrons. Originally the Hubbard model was studied intending applications to the metal–insulator transition, and to the problem of itinerant magnetism in narrow bands in metallic crystals. Then it has become a very popular starting point (in its two-dimensional version) to describe the electrons in the Cu_2O-planes of high-temperature superconducting materials. Although it has been studied very intensively with a variety of methods, important aspects of its properties are still incompletely understood, and thus the model (5.58) still is an active area of research.

For higher-dimensional systems, the transformation of (5.55) does not help, and thus a different approach for dealing with the fermion operators is needed. One strategy is to integrate out the fermionic degrees of freedom by introducing auxiliary bosonic fields. For this purpose one uses the identity

$$\int_{-\infty}^{+\infty} \exp(-a\phi^2 - b\phi) \mathrm{d}\phi = \sqrt{\pi/a} \exp(b^2/4a), \tag{5.59}$$

where a, b are real numbers, and $a > 0$ is required. Suppose now we wish to calculate the grand-canonical partition function Z_{gc} with the Trotter–Suzuki formula,

$$\begin{aligned} Z_{\mathrm{gc}} &= \mathrm{Tr} \exp\left[-\frac{1}{k_{\mathrm{B}} T} \left(\hat{\mathcal{H}} - \mu \hat{\mathcal{N}} \right) \right] = \lim_{P \to \infty} Z^P_{\mathrm{gc}}, \\ Z^P_{\mathrm{gc}} &= \mathrm{Tr} \left\{ \left(\exp\left[-\frac{1}{k_{\mathrm{B}} T P} \left(\hat{E}^{\mathrm{kin}} - \mu \hat{\mathcal{N}} \right) \right] \exp\left(-\frac{1}{k_{\mathrm{B}} T P} \hat{V} \right) \right)^P \right\}, \end{aligned} \tag{5.60}$$

μ being the chemical potential, and we have made use of the Trotter formula to disentangle terms which are quadratic in the fermion operators ($\hat{E}^{\mathrm{kin}} - \mu \hat{\mathcal{N}}$) from quartic terms ($\hat{V}$ thus contains the terms $U \hat{n}_{i\uparrow} \hat{n}_{i\downarrow} = U \hat{c}^+_{i\uparrow} \hat{c}_{i\uparrow} \hat{c}^+_{i\downarrow} \hat{c}_{i\downarrow}$). Now one can express $\exp(-\hat{V}/k_{\mathrm{B}} T P)$ as an exponent of a quadratic form, if one uses (5.59) as follows ($U > 0$)

$$\exp\left[-\frac{U}{k_{\mathrm{B}}TP}\sum_{i=1}^{N}\hat{n}_{i\uparrow}\hat{n}_{i\downarrow}\right]$$
$$\propto\prod_{i=1}^{N}\int_{-\infty}^{+\infty}\mathrm{d}\phi_i\exp\left[-\frac{k_{\mathrm{B}}TP\phi_i^2}{2U}-\phi_i(\hat{n}_{i\uparrow}-\hat{n}_{i\downarrow})-\frac{U(\hat{n}_{i\uparrow}+\hat{n}_{i\downarrow})}{2k_{\mathrm{B}}TP}\right]. \tag{5.61}$$

Symbolically, the Pth approximant $\mathcal{Z}_{\mathrm{gc}}^{P}$ to the grand-canonical partition function can be written as follows, involving a quadratic form of the fermion operators in the exponential,

$$\mathcal{Z}_{\mathrm{gc}}^{P}=\prod_{i=1}^{N}\int_{-\infty}^{+\infty}\mathrm{d}\phi_i\mathrm{Tr}\left\{\exp\left[\hat{c}_{i\sigma}^{+}A_{ij}^{(1)}\hat{c}_{j\sigma}\right]\cdots\exp\left[\hat{c}_{i\sigma}^{+}A_{ij}^{(2P)}\hat{c}_{j\sigma}\right]\right\}, \tag{5.62}$$

where the $A_{ij}^{(k)}$ (which depend on the ϕ_i and σ, of course) are elements of $N\times N$-matrices $\underline{A}^{(k)}$. However, for quadratic forms in the fermion operators as appear in (5.62) the trace over the fermionic degrees of freedom can be carried out analytically, to yield [5.18]

$$\mathrm{Tr}\left\{\exp\left[\hat{c}_i^{+}\hat{A}_{ij}^{(1)}\hat{c}_j\right]\cdots\exp\left[\hat{c}_i^{+}\hat{A}_{ij}^{(P)}\hat{c}_j\right]\right\}=\det\left\{1+\exp\left(\underline{A}^{(1)}\right)\cdots\exp\left(\underline{A}^{(P)}\right)\right\}. \tag{5.63}$$

As a result, the partition function becomes [5.18]

$$\begin{aligned}Z_{\mathrm{gc}}^{(P)}\propto\prod_{i,s}\int\mathrm{d}\phi^{(s)}&\exp\left[-\frac{k_{\mathrm{B}}TP\left(\phi_i^{(s)}\right)^2}{2U}\right]\\&\times\det\left[\underline{1}+\exp(-\widetilde{K}/k_{\mathrm{B}}TP)\exp\left(-\widetilde{V}\left(\phi_i^{(1)}\right)\right)\right.\\&\left.\times\exp(-\widetilde{K}/k_{\mathrm{B}}TP)\cdots\exp\left(-\widetilde{K}/k_{\mathrm{B}}TP\right)\exp\left(-\widetilde{V}\left(\phi_i^{(P)}\right)\right)\right]\\&\times\det\left[\underline{1}+\exp(-\widetilde{K}/k_{\mathrm{B}}TP)\exp\left(-\widetilde{V}\left(-\phi_i^{(1)}\right)\right)\exp(-\widetilde{K}/k_{\mathrm{B}}TP)\right.\\&\qquad\left.\cdots\exp(-\widetilde{K}/k_{\mathrm{B}}TP)\exp\left(-\widetilde{V}\left(-\phi_i^{(P)}\right)\right)\right].\end{aligned} \tag{5.64}$$

Here $\widetilde{K}$ is an abbreviation for $\widetilde{K}=E_{\mathrm{kin}}-(\mu-V/2)\mathcal{N}$, the kinetic energy matrix for a single-particle matrix on a lattice, and $\widetilde{V}(\phi_\ell)$ is a diagonal matrix depending on the $\phi_{\ell j}$ which we do not write out in detail here.

By eliminating the fermionic degrees in favor of Gaussian fields $\phi_i^{(s)}$ (or in favor of Ising spin variables, which is an even more useful alternative for the Hubbard model [5.26]), one has managed to express the partition function entirely in terms of real numbers, so it looks suitable for a Monte Carlo evaluation. However, in order to be able to interpret the result in terms of an effective classical Hamiltonian, i.e.,

$$Z_{\mathrm{gc}}^{(P)} \propto \prod_{i,s} \int \mathrm{d}\phi_i^{(s)} \exp\left[-\mathcal{H}_{\mathrm{eff}}^{(P)}\left(\phi_i^{(s)}\right)/k_{\mathrm{B}}T\right], \tag{5.65}$$

it is mandatory that the integrand in (5.65) is always non-negative, and this is not the case! This problem – which is quite typical for quantum Monte Carlo simulations of fermionic systems – is called the "minus-sign problem" [5.27]. So when we want to calculate an average of a quantity $A(x)$ with a measure $\varrho(x)$,

$$\langle A\rangle = \int A(x)\varrho(x)\mathrm{d}x \Big/ \int \varrho(x)\mathrm{d}x, \tag{5.66}$$

we can no longer interpret $\varrho(x)$ as probability density, if $\varrho(x)$ is not positive semi-definite, and so the basis for Metropolis importance sampling is no longer valid. Of course, this difficulty can be circumvented formally by the trick of using the probability density $\widetilde{\varrho} = |\varphi(x)|/\int|\varrho(x)|\mathrm{d}x$ and absorbing the sign of $\varrho(x)$ in the quantity that one wishes to estimate,

$$\langle A\rangle = \frac{\int A(x)\mathrm{sign}(\varrho(x))\widetilde{\varrho}(x)\mathrm{d}x}{\int \mathrm{sign}(\varrho(x))\widetilde{\varrho}(x)\mathrm{d}x} = \frac{\langle A\hat{s}\rangle}{\langle\hat{s}\rangle}, \tag{5.67}$$

where $\hat{s}$ is the sign operator. But it is unlikely that this importance sampling based on $|\varrho(x)|$ will really sample the important regions of phase space when N gets large. Indeed, for the Hubbard Hamiltonian one estimates that [5.18]

$$\langle\hat{s}\rangle \propto \exp(-\gamma NU/k_{\mathrm{B}}T), \tag{5.68}$$

where γ is a constant of order unity. Thus it is clear that for low temperatures and large N the denominator $\langle\hat{s}\rangle$ in (5.67) gets extremely small, and hence it becomes impossible to sample $\langle A\rangle$ using (5.67) with meaningful accuracy.

While many schemes have been devised to alleviate this problem, a fully satisfactory solution to this "minus sign problem" is unknown to the authors of this book at the time of writing. In view of these difficulties, we have confined ourselves to a rather sketchy description of the quantum Monte Carlo approach to fermions on lattices, since this is still an active area of research. Also the treatment of quantum spin models still is under development: in particular, substantial improvement has been obtained by combining cluster algorithms with PIMC Monte Carlo methods for quantum spin systems [5.28].

5.4 Concluding Remarks

Quantum Monte Carlo simulation is a particularly rich field, and many aspects are still under development. In this chapter, we have emphasized the path integral quantum Monte Carlo technique and even for this method only the flavor of the approach could be given, and typical applications were sketched. There are also important Monte Carlo methods addressing the problem of solving the many-body Schrödinger equation, in order to find ground-state energy and associated wave functions [5.18]. We refer the interested reader to the literature where he can find very concise accounts of "variational Monte Carlo (VMC)", "Green's function Monte Carlo (GFMC)", "Projector Quantum Monte Carlo (PQMC)", etc., [5.18]. Path integral simulations of "rotors" (rotating rigid molecules) have been reviewed in [5.29].

Chapter 6
Monte Carlo Methods for the Sampling of Free Energy Landscapes

6.1 Introduction and Overview

In this chapter, we return to classical statistical mechanics. In the canonical ensemble averages of an observable $A(\vec{x})$, where $\vec{x}$ stands symbolically for the "microstate" coordinate in the configurational part of the phase space of the system, are given by (cf. Sect. 2.1.1)

$$\langle A(\vec{x})\rangle_T = \frac{1}{Z}\int \mathrm{d}\vec{x}\,\exp\left[-\mathcal{H}(\vec{x})/k_\mathrm{B}T\right]A(\vec{x}). \tag{6.1}$$

Here $\mathcal{H}(\vec{x})$ is the Hamiltonian of the considered many-particle system, k_B Boltzmann's constant, T absolute temperature, and Z being the partition function, which is related to the free energy F,

$$Z = \int \mathrm{d}\vec{x}\,\exp[-\mathcal{H}(\vec{x})/k_\mathrm{B}T], \quad F = -k_\mathrm{B}T\ln Z. \tag{6.2}$$

By simple sampling Monte Carlo (Sects. 2.1.2–2.1.5), the integrals over the (very high-dimensional!) space $\int \mathrm{d}\vec{x}\ldots$in (6.1), (6.2) are replaced by sums over a sample of M randomly chosen points $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_M\}$, but we have seen that simple sampling works only if the number N of the particles (or degrees of freedom, such as Ising or Heisenberg spins, cf. Sect. 2.1.1) is extremely small. Therefore, we have introduced the Importance Sampling Monte Carlo method (Sect. 2.1.6), where the points $\{\vec{x}_\nu\}$ no longer are chosen completely at random, but preferentially from the important region of configuration space. Namely, choosing a state $\vec{x}_\nu$ with a probability proportional to the Boltzmann factor $\exp(-\mathcal{H}(\vec{x}_\nu)/k_\mathrm{B}T)$, the average in (6.1) is simply replaced by an arithmetic average over the M states generated, cf. (2.35):

$$\langle A(\vec{x})\rangle_T \approx \overline{A(\vec{x})} = \frac{1}{M}\sum_{i=1}^{M} A(\vec{x}_i). \tag{6.3}$$

K. Binder and D. W. Heermann, *Monte Carlo Simulation in Statistical Physics*,
Graduate Texts in Physics, https://doi.org/10.1007/978-3-030-10758-1_6

In the previous chapters of this book, we have seen that the Importance Sampling Monte Carlo method is very powerful; it allows a very large number of useful applications. However, when we compare (6.1) and (6.3), we see that some important information has been lost: (6.3) has as a normalizing denominator no longer the partition function Z (6.2), but rather simply the total number M of generated configurations. In fact, doing Importance Sampling Monte Carlo, the knowledge of Z and hence F, as well as of the entropy S (remember $F = E - TS$, where the internal energy E is accessible as a thermal average of the Hamiltonian, of course, $E = \langle \mathcal{H}(\vec{x}\rangle_T)$ has been lost.

This lack of knowledge on F is particularly disadvantageous when one deals with first-order phase transitions, of course, since there several "macrostates" compete, which at the transition all have the same free energy. Consider, for example, the thermally driven phase transition of the q-state Potts ferromagnet, that was already considered in Sect. 2.3.7 in a finite size scaling context. When we study (in the thermodynamic limit, $N \to \infty$) the free energy per spin F/N of the Potts ferromagnet as a function of temperature (in zero field H), we observe that for temperatures T less than the transition temperature T_c the free energy of the (q-fold degenerate) ordered phase is lower than that of the (metastable) continuation of the free energy branch of the (nondegenerate) disordered phase, while for $T > T_c$ the disordered branch has the lower free energy. At T_c, the two branches $F_{\mathrm{ord}}(T)$, $F_{\mathrm{dis}}(T)$ are precisely equal, but the two branches meet there under different slope. Since we know from elementary thermodynamics that ($\beta \equiv 1/k_B T$)

$$E = (\partial F/\partial\beta)_H = -k_B T^2 \left(\frac{\partial F}{\partial T}\right)_H , \tag{6.4}$$

the slopes $\partial F_{\mathrm{ord}}/\partial T|_{T_c}$, $\partial F_{\mathrm{dis}}/\partial T|_{T_c}$ simply are related to the energies E_-, E_+ in Fig. 2.20, $E_+ - E_-$ being the latent heat at the first-order transition. Since it turns out that the phases that coexist at the first-order transition are separated by a free energy barrier (that is huge if N is sufficiently large), the finite size scaling analysis of first-order phase transitions (described in Sect. 2.3.7) is often difficult to apply, when one relies on straightforward importance sampling. So, one would like to know the free energies of the ordered and disordered phases $F_{\mathrm{ord}}(T)$, $F_{\mathrm{dis}}(T)$ explicitly, since then the transition simply could be located from the condition $F_{\mathrm{ord}}(T_c) = F_{\mathrm{dis}}(T_c)$. The "poor man's recipe" to achieve this has been based on a method called "thermodynamic integration" [6.1, 6.2]. From (6.4), we readily recognize that free energy differences can be computed from

$$\Delta F \equiv F_2 - F_1 = \int_{\beta_1}^{\beta_2} E(\beta)\mathrm{d}\beta = \int_{\beta_1}^{\beta_2} \langle \mathcal{H}\rangle_\beta \mathrm{d}\beta. \tag{6.5}$$

To carry out the integral in (6.5) numerically with sufficient accuracy, one needs to compute $\langle \mathcal{H}\rangle_\beta$ at a large number of inverse temperatures β intermediate between β_1 and β_2 (and this need may be cumbersome: that is why we call this approach a

"poor man's recipe"). To get an absolute free energy, one needs to know the free energy F_1 of a reference state. For the Potts ferromagnet in the disordered phase, a convenient reference state is the completely random state at $\beta_1 = 0$, where $E(0) = 0$ and $S(\beta = 0) = Nk_B \ln q$ is trivially known. For the ordered phase, a similarly simple reference state would be $\beta_2 \to \infty$, where the system is fully ordered, E coincides with the ground state energy, and $S = k_B \ln q$ (without a factor N: we just have a q-fold degenerate ground state). In practice, it suffices to choose a large but finite β_2 as a reference state, for which the deviation of $E(\beta_2)$ from the ground state energy is negligibly small.

Although we have called this thermodynamic integration approach a "poor man's recipe", one should not be misled: There are cases where this simple-minded approach actually is the method of choice, superior in accuracy to both finite size scaling approaches and the sampling techniques that we shall describe in this chapter; examples for the high accuracy that one can achieve in the location of transition temperatures of lattice models can be found in [6.3–6.5]. Some of these examples concern surface effects at first-order phase transitions ("surface-induced ordering", "surface-induced disordering" [6.6]), and in such context one needs to locate the transition temperature of a very large model system in the bulk with extremely high accuracy to allow for meaningful conclusions.

Even though thermodynamic integration can (at least in favorable cases) yield information on the free energies of the two phases that coexist at the transition, it does not yield any information on the barrier between these phases in configuration space. This free energy barrier is responsible for the hysteresis that very often is observed in simulations and also in experiments!) near first-order transitions, cf. Fig. 2.20. In many contexts, it is of great interest to gain information on a suitable path in a (suitably coarse-grained) "free energy landscape" that connects the free energy minima that correspond to the coexisting phases with each other.

When one discusses $\mathcal{H}(\vec{x})$ as a function of the configuration space coordinate $\vec{x}$ one may speak of a "potential energy landscape", but clearly the considered space is extremely high-dimensional, and the landscape in this space is presumably extremely "rugged". As always in statistical physics, one is more interested in a reduced description, referred to as "coarse-graining" above. The most familiar concept of coarse-graining uses the idea of an "order parameter" as introduced by Landau [6.1]. In simple cases, like Ising ferromagnets or fluids that undergo a vapor–liquid phase transition, this order parameter is a scalar quantity (the magnetization $\langle m \rangle_{T,H}$ per spin in the case of a ferromagnet, or the density $\rho = \langle N \rangle_T / V$ in the case of a fluid occupying a volume V). Rather than discussing the full canonic or grand-canonic probability distribution

$$p_{\text{can}}(\vec{x}) = (1/Z) \exp[-\mathcal{H}(\vec{x})/k_B T]$$

or

$$p_{g,c}(\vec{x}) = (1/Y) \exp[(\mu N - \mathcal{H}(\vec{x}))/k_B T], \tag{6.6}$$

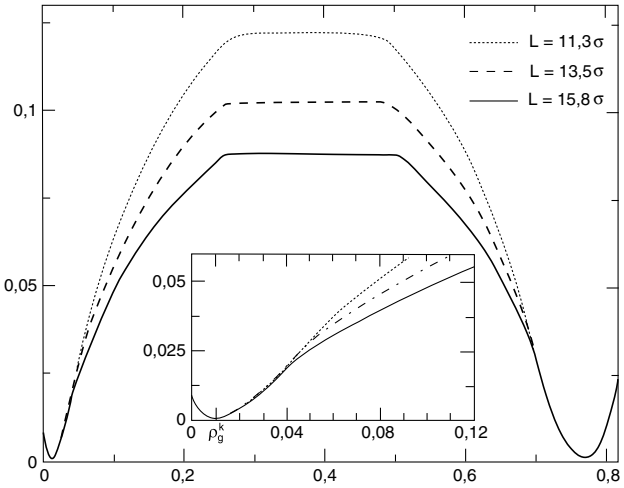


Fig. 6.1 Free energy per unit volume $f(\rho)$ plotted versus density ρ, for a Lennard–Jones model of a fluid (the potential being truncated at $r_c = 2.2^{1/6}\sigma$, with the Lennard–Jones parameter σ being the unit of length, and shifted to zero there). The data refer to three choices of cubic $L \times L \times L$ simulation boxes, with periodic boundary conditions, a temperature $T = 0.68T_c$, and the chemical potential being chosen at two-phase coexistence, $\mu = \mu_{coex}$, so the two minima are equally deep. *Insert* shows a blow-up of the region near the bulk gas density at coexistence, ρ_g

where $Y(T, V, \mu) = \sum_N \exp(\mu N/k_B T) Z_N(T, V)$ is the grand-canonical partition function and μ the chemical potential, one focusses attention to a reduced description, for example, in terms of the distribution $p(\rho)$ of the density only,

$$p(\rho) = (1/Y) \sum_N \int d\vec{x} \exp[(\mu N - \mathcal{H}(\vec{x}))/k_B T] \delta(\rho - N/V). \tag{6.7}$$

Note that $p(\rho)$ can be interpreted in terms of a free energy $f(\rho)$ per unit volume as

$$p(\rho) = (1/Y) \exp[-V f(\rho)/k_B T]. \tag{6.8}$$

Figure 6.1 presents a plot of $f(\rho)$ versus ρ for a Lennard–Jones model of a fluid at a temperature $T = 0.68T_c$ and choosing the chemical potential μ such that in the bulk (volume $V = L^3 \to \infty$) coexistence between vapor and liquid occurs, $\mu = \mu_{coex}(T)$ [6.7, 6.8]. These data have been obtained by the technique of "successive umbrella sampling" [6.9], which will be described in the following section. Related data can be found in [6.10–6.12] for fluids and in [6.13–6.15] for Ising (lattice gas) models. One can see that the graph of $f(\rho)$ versus ρ exhibits several (rounded) kinks; this behavior (which leads to true singularities in the thermodynamic limit, $L \to \infty$) is more clearly visible when one considers the variation of $\beta\hat{\mu}$ with ρ, where $\mu = (\partial f/\partial \rho)_T$ and $\hat{\mu} = \mu - \mu_{coex}(T)$, see Fig. 6.2 [6.7, 6.8]. The small snapshot pictures under-

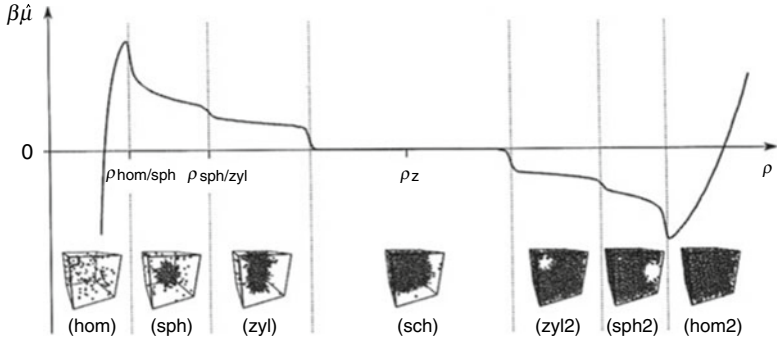


Fig. 6.2 Plot of the reduced chemical potential difference $\beta\hat{\mu}$ versus ρ, for the system with $L = 15.8\sigma$ from Fig. 6.1. Typical snapshot pictures of system configurations illustrate the evolution of the system from homogeneous vapor to homogeneous liquid via a sequence of inhomogeneous states when one crosses the free energy barrier of Fig. 6.1 when ρ is increased. For further explanations cf. text

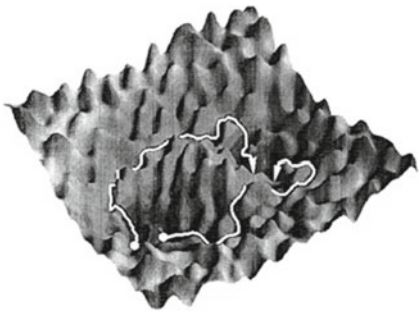


Fig. 6.3 Schematic picture of a "rugged free energy landscape", plotted as a function of two order parameter coordinates ψ_x, ψ_y. The two *white lines* indicate two possible trajectories from the region of one deep minimum to another one. From Dellago et al. [6.28]

neath this curve illustrate the different states of the system: For $\rho < \rho_{\text{hom/sph}}$ the vapor phase is homogeneous, while near $\rho = \rho_{\text{hom/sph}}$ a transition takes place to an inhomogeneous state of the system, where a spherical droplet coexists with surrounding (supersaturated) vapor. The so-called "droplet evaporation–condensation transition" has received a lot of attention both by simulations [6.7, 6.8, 6.10–6.14] and by theory [6.16–6.19] (Fig. 6.2).

From an analysis of the free energy barrier in Fig. 6.1 in the region $\rho_{\text{hom/sph}} < \rho < \rho_{\text{sph/cyl}}$, one can extract valuable information on the surface free energy of nanoscopic droplets [6.7, 6.8, 6.10, 6.13, 6.16, 6.20]. Then, for $\rho_{\text{sph/cyl}} < \rho < \rho_{\text{cyl/slab}}$ a regime occurs, where the periodic boundary conditions stabilize a cylindrical liquid droplet rather than a spherical one, before for $\rho_{\text{cyl/slab}} < \rho < \rho_{\text{slab/cyl}}$ the periodic boundary conditions stabilize a liquid slab. Varying the density does not change the total surface area (which is $2L^2$) in this region, only the amount of liquid phase changes. Since the two liquid–vapor interfaces occurring then in the system do not have any curvature on average, one has $\hat{\mu} = 0$ in this region, $P(\rho)$ is a flat function (independent of ρ, if mutual interactions between the two interfaces are negligible). The free

energy difference Δf between this state and the homogeneous states at two-phase coexistence thus simply is $\Delta f = 2L^2 f_{\text{int}}$, where f_{int} is the interfacial tension. Hence, measurement of Δf has been proposed [6.21] as a method to estimate f_{int} via

$$f_{\text{int}} = \lim_{L\to\infty} (1/2L^2)\Delta f. \tag{6.9}$$

Indeed this technique has found widespread and successful use (e.g., [6.14, 6.15, 6.22–6.27).

For $\rho_{\text{slab/cyl}} < \rho < \rho_{\text{cyl/bub}}$, one has again a cylindrical inhomogeneity stabilized by the periodic boundary condition, but now the role of vapor and liquid are interchanged, we find a cylinder of vapor surrounded by liquid and, for $\rho_{\text{cyl/bub}} < \rho < \rho_{\text{bub/hom}}$, a spherical vapor bubble surrounded by liquid. Again, an analysis of the free energy $f(\rho)$ in this region yields interesting information on the surface free energy of bubbles and free energy barriers for bubble nucleation. Thus, the path over the free energy "mountain" that one has to take when one moves from homogeneous vapor to homogeneous liquid by increasing the density at fixed volume of the system contains a lot of physically relevant information, elucidating interfacial phenomena and the kinetics of transitions from one stable phase to another.

Of course, we observe only a free energy "mountain" since in the step from (6.6) to (6.7) we have drastically reduced the problem by integrating over all degrees of freedom apart from a single scalar variable, the density ρ is this case, considering then $f(\rho)$. When we consider a coarse grained free energy as a function of several variables (which we formally combine into a vector $\vec{\Psi}$), these minima are not close-by located with respect to each other, multiple trajectories may need consideration, in particular if the free energy "landscape" is a rugged landscape of hills and valleys (Fig. 6.3) [6.28]. Such "rugged free energy landscapes" [6.29] occur in such diverse systems as domain patterns in random ferromagnets, diluted magnets with competing interactions ("spin glasses" [6.30], undercooled fluids near the glass transition [6.31], and polymers (in particular proteins) under bad solvent conditions [6.32–6.40]. In all these systems, the step analogous to the step from (6.6) to (6.8) for the simple fluid, where one focuses on a reduced description in terms of a simple "order parameter", is not at all straightforward. In models for spin glasses, this problem has been circumvented by studying two "real replica", copies of the same system with a specific choice of random exchange interactions, and defining an order parameter q as the (normalized) projection of the spin configurations in the replicas (1) and (2) onto each other, $q = (1/N)\sum_{i=1}^{N} S_i^{(1)} S_i^{(2)}$ (in the case of Ising spins, $S_i^{(\alpha)} = \pm 1$) [6.30, 6.31]. The sampling of the distribution $P(q)$, analogous to the distribution $P(\rho)$ in (6.7) and (6.8), nevertheless is a challenging computational problem, since in addition to the thermal averaging (with the Boltzmann factor, cf. (6.7)) there is the need to carry out an average $[\cdots]_{\text{av}}$ over the quenched disorder: that is, the distribution $P_{\{J_{ij}\}}(q)$ obtained for one particular choice $\{J_{ij}\}$ of the random bonds needs to be averaged over (typically) several hundred realizations of such random bond configurations [6.30, 6.41–6.46]. For the structural glass transition of undercooled fluids and for protein folding, on the other hand, even an understanding

what order parameters one should use to describe these problems is lacking [6.31, 6.34]. Despite the great interest in these "grand challenge problems" of physics, we shall not emphasize them in this chapter.

However, in all these problems with rugged free energy landscapes, a useful first step involves the consideration of the energy density of states $g(E)$. This quantity is introduced rewriting the partition function in (6.2) as

$$Z = \int \mathrm{d}\vec{x} \exp[-\mathcal{H}(\vec{x})/k_\mathrm{B}T] = \int \mathrm{d}E g(E) \exp(-E/k_\mathrm{B}T). \tag{6.10}$$

In full analogy with (6.7), it can be defined as

$$g(E) = \frac{1}{Z} \int \mathrm{d}\vec{x} \exp[-\mathcal{H}(\vec{x})/k_\mathrm{B}T] \delta[E - \mathcal{H}(\vec{x})]. \tag{6.11}$$

Sometimes, it is useful to generalize (6.11) to make it a function of several variables: For example, when one considers the problem of adsorption of an end-grafted flexible macromolecule (a so-called "polymer mushroom") on an attractive substrate surface under bad solvent conditions [6.39], it is useful to consider instead of $g(E)$ the function $g(E_\mathrm{s}, E_\mathrm{int})$, where $E = E_\mathrm{s} + E_\mathrm{int}$, E_s being the energy of adsorption won by the monomers at the attractive surface, while E_int is the interaction energy between monomers. Such distributions $g(E)$ or generalizations therefore can be sampled by various extensions of "umbrella sampling" [6.47, 6.48] such as "multicanonical Monte Carlo" [6.49–6.55] (Sect. 6.3) or the Wang–Landau method [6.56–6.63] (Sect. 6.4), for instance.

The implementation of these "extended ensemble methods" [6.64, 6.65] and related methods [6.65–6.74] has been a very active area of research during the last two decades, and hence we give here only a simple introduction to some of the concepts that have found the most widespread use. As an example for the usefulness of such methods, Fig. 6.4 shows the "phase portrait" of a tethered polymer chain described by the bond-fluctuation model on the simple cubic lattice [6.75, 6.76], a system [6.39] already mentioned above, where $E_\mathrm{s} = \sum_i \varepsilon_\mathrm{s} n_\mathrm{s}$, n_s being the number of monomers that are in contact with the substrate surface and win an energy ε_s each, $\beta_\mathrm{s} = \varepsilon_\mathrm{s}/k_\mathrm{B}T$, and $E_\mathrm{int} = \sum_{\langle i,j \rangle} \varepsilon_\mathrm{b} n_{ij}$, n_{ij} being the number of bead–bead pairs that win an interaction energy ε_b each, and the sums over i, j run over all the monomers, $\beta_\mathrm{b} = \varepsilon_\mathrm{b}/k_\mathrm{B}T$. Due to the competition between the two energy scales and the constraints (connectivity of the chain, anchoring at the surface), this macromolecule model exists in many different conformations (note that the layered states are, at least in part, crystalline states), and without the Wang–Landau algorithm [6.56–6.58] it would have required an enormous effort to map out the behavior of the model in the $(\beta_\mathrm{b}, \beta_\mathrm{s})$ plane (Fig. 6.4) in detail.

Of course, a "phase portrait" as shown in Fig. 6.4 tells us only which states dominate for a given point $(\beta_\mathrm{s}, \beta_\mathrm{b})$ in the parameter space; it does not tell us anything about the path that the system takes when we start in a state distinct from equilib-

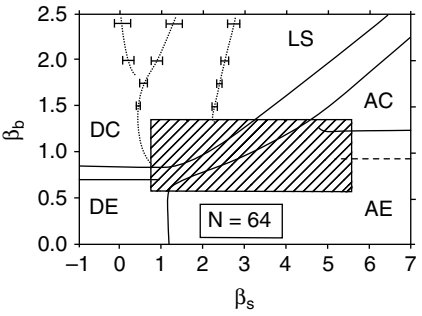


Fig. 6.4 Phase portrait for a tethered polymer chain of length $N = 64$ in the space of the surface coupling (β_s) and bead–bead coupling (β_b). The *solid lines* represent maxima of surface energy and bead–bead contact energy fluctuations, and separate different states of the "polymer mushroom": Desorbed expanded (*DE*), adsorbed expanded (*AE*), desorbed collapsed (*DC*), adsorbed collapsed (*AC*) and layered structures (*LS*). The *dashed line* is an estimate for the location of the coil–globule transition, while the *dotted lines* represent shallow maxima of the surface energy fluctuation. In the *shaded area* the free energy landscape could not be resolved well enough to identify states that are well separated from each other

rium, in order to approach equilibrium as a function of "time" (recall the dynamic interpretation of Monte Carlo sampling, see Chap. 2).

Given a rugged free energy landscape as schematically shown in Fig. 6.3, it is a nontrivial question which path the system will take to reach equilibrium. To describe the kinetics of such thermally activated processes in complex free energy landscapes, a Monte Carlo sampling over an ensemble of trajectories of such processes is desirable, and this can be achieved by the "transition path sampling" technique [6.28, 6.77–6.83] and its variants to sample rare events ("string method", "metadynamics", etc.) [6.84–6.90]. As Bolhuis et al. [6.80] put it, transition path sampling can be viewed as "throwing ropes over mountain passes in the dark".

Of course, the sampling of trajectories of processes requiring thermal activation such as crossing barriers in nucleation events via Monte Carlo methods has its sound theoretical basis in the dynamic interpretation of Monte Carlo sampling [6.91]. In fact, the first Monte Carlo simulation studies of nucleation processes in the nearest-neighbor kinetic Ising model were performed decades ago [6.92, 6.93]. However, this naive (i.e., unbiased) sampling of trajectories over free energy barriers is efficient only if these barriers are not higher than a few $k_B T$. Via transition path sampling, one creates trajectories that are biased such that higher free energy barriers (of the order of 20–100 $k_B T$) are crossed, and thus very valuable insight into the kinetics of nucleation processes (e.g., [6.94]) can be gained. In Sect. 6.5, we shall give the flavor of this transition path sampling method (more details can be found in basic reviews [6.28, 6.80, 6.83]).

6.2 Umbrella Sampling

We now return to the problem of calculating free energy differences between two states 1 and 2, but we now assume these systems have different Hamiltonians $\mathcal{H}_1(\vec{x})$, $\mathcal{H}_2(\vec{x})$. From (6.2), we can immediately derive that

$$\Delta F = F_2 - F_1 = -k_B T \ln(Z_2/Z_1) = -k_B T \ln\left(\frac{\int \mathrm{d}\vec{x} \exp[-\beta\mathcal{H}_2(\vec{x})]}{\int \mathrm{d}\vec{x} \exp[-\beta\mathcal{H}_1(\vec{x})]}\right). \quad (6.12)$$

Following the very pedagogic account given by Frenkel and Smit [6.95], we first use (6.12) to justify the "overlapping distribution method" [6.96]. Suppose that we carry out an importance sampling study of system 2. For every configuration space point $\vec{x}$ that is generated in the resulting Markov chain, we can compute also the potential energy $\mathcal{H}_1(\vec{x})$ of system 1. (We focus here on the case that both systems have exactly the same volume V and same particle number N.) We then can obtain the probability of $\Delta E = \mathcal{H}_2(\vec{x}) - \mathcal{H}_1(\vec{x})$ as

$$p_2(\Delta E) = (1/Z_2) \int \mathrm{d}\vec{x} \exp[-\beta\mathcal{H}_2(\vec{x})]\delta[\mathcal{H}_2(\vec{x}) - \mathcal{H}_1(\vec{x}) - \Delta E]. \quad (6.13)$$

Due to the delta function, (6.13) can be rewritten as

$$\begin{aligned} p_2(\Delta E) &= (1/Z_2)\exp(-\beta\Delta E) \int \mathrm{d}\vec{x} \exp[-\beta\mathcal{H}_1(\vec{x})]\delta[\mathcal{H}_2(\vec{x}) - \mathcal{H}_1(\vec{x}) - \Delta E] \\ &= (Z_1/Z_2)\exp(-\beta\Delta E)\, p_1(\Delta E) \end{aligned} \quad (6.14)$$

In full analogy to (6.13), $p_1(\Delta E)$ is the probability density of finding a potential energy density ΔE between systems 1 and 2, when one samples the states of the system 1. Using (6.12), we obtain

$$\ln p_2(\Delta E) = \beta(\Delta F - \Delta E) + \ln p_1(\Delta E). \quad (6.15)$$

From (6.15), we see that the two functions $\ln p_2(\Delta E)$ and $\ln p_1(\Delta E) - \beta\Delta E$ should be identical, apart from a constant offset $\beta\Delta F$. Thus, if there exists a range of ΔE where the two distributions $\ln p_2(\Delta E)$ and $\ln p_1(\Delta E) - \beta\Delta E$ have sufficient overlap, one can obtain $\beta\Delta F$ from a best fit. Of course, in practice, this method can work only if the two systems differ only very little (e.g., two Ising systems at slightly different values of the magnetic field H).

Another consequence of (6.14) is seen when one integrates this equation over ΔE from $-\infty$ to $+\infty$. Since p_2 is normalized to unity, this yields

$$1 = \exp(\beta\Delta F) \int_{-\infty}^{+\infty} \mathrm{d}\Delta E \exp(-\beta\Delta E)\, p_1(\Delta E) = \exp(\beta\Delta F)\langle\exp(-\beta\Delta E)\rangle_1,$$

that is

$$\exp(-\beta\Delta F) = \langle \exp\{-\beta(\mathcal{H}_2(\vec{x}) - \mathcal{H}_1(\vec{x}))\}\rangle_1. \tag{6.16}$$

Here $\langle \cdots \rangle_1$ means that averages are taken in a sampling with system 1. Again (6.16) is practically useful if the two systems 1 and 2 differ only very little.

The idea of "umbrella sampling" [6.47, 6.48] now is to enhance the overlap between the distributions $p_1(\Delta E)\exp(-\beta\Delta E)$ and $p_2(\Delta E)$ by carrying out a biased sampling with a suitable chosen weight function $\Pi(\vec{x})$. Biased sampling is derived by rewriting (6.1) as follows

$$\begin{aligned}\langle A(\vec{x})\rangle_0 &= \frac{\int \mathrm{d}\vec{x}\,\Pi(\vec{x})\exp[-\beta\mathcal{H}_0(\vec{x})]A(\vec{x})/\Pi(\vec{x})}{\int \mathrm{d}\vec{x}\,\Pi(\vec{x})\exp[-\beta\mathcal{H}_0(\vec{x})]/\,\Pi(\vec{x})} \\ &= \langle A(\vec{x})/\Pi(x)\rangle_{0,\Pi}/\langle 1/\Pi(\vec{x})\rangle_{0,\Pi},\end{aligned} \tag{6.17}$$

where the index Π at $\langle \cdots \rangle_{0,\Pi}$ is a reminder that the weight contains an additional factor $\Pi(\vec{x})$ in the statistical weight for the states $\vec{x}$, in addition to the Boltzmann factor.

We use this to rewrite (6.16) as follows

$$\begin{aligned}&\langle \exp\{-\beta[\mathcal{H}_2(\vec{x}) - \mathcal{H}_1(\vec{x})]\}\rangle_1 \\ &= \frac{\int \mathrm{d}\vec{x}\,\Pi(\vec{x})\exp[-\beta\mathcal{H}_2(\vec{x})]/\Pi(\vec{x})}{\int \mathrm{d}\vec{x}\,\Pi(\vec{x})\exp[-\beta\mathcal{H}_1(\vec{x})]/\Pi(\vec{x})} \\ &= \langle \exp[-\beta\mathcal{H}_2(\vec{x})]/\Pi(\vec{x})\rangle_\Pi/\langle \exp[-\beta\mathcal{H}_1(\vec{x})]/\Pi(\vec{x})]\rangle_\Pi.\end{aligned} \tag{6.18}$$

Here, (6.17) is used for the special case $\mathcal{H}_0(\vec{x}) = 0$. In order to be able to sample both the numerator and the denominator with acceptable accuracy, the "bridging distribution" $\Pi(\vec{x})$ must have sufficient overlap with both the important regions of configuration space of systems 1 and 2. This bridging property of $\Pi(\vec{x})$ is responsible for the name umbrella sampling.

In practice, it is often advantageous to work not only with a single bridging distribution $\Pi(\vec{x})$, but with many distributions intermediate between $\frac{1}{Z_1}\exp[-\beta\mathcal{H}_1(\vec{x})]$ and $\frac{1}{Z_2}\exp[-\beta\mathcal{H}_2(\vec{x})]$. Suppose, there exists a parameter $\Delta\psi$ that distinguishes between $\mathcal{H}_2$ and $\mathcal{H}_1$. Then, it often is useful to divide the interval $\Delta\psi$ in n steps $\Delta\psi/n$. The optimum choice of n, however, in the general case is a subtle matter [6.95]. Another important drawback is that the function $\Pi(\vec{x})$ is not known a priori, it must be guessed, using the knowledge on the Boltzmann weights of the systems 1 and 2. As Frenkel and Smith [6.95] put it, "constructing a good sampling distribution used to require skill and patience", and therefore initially this method did not find widespread use. However, there are situations where the implementation of the method is rather straightforward. One such scheme is the technique of Virnau and Müller [6.9] termed "successive umbrella sampling", to which we turn now. This method has the advantage that by extrapolation of results from one "window" to the next (adjacent) window one can estimate the weight function rather simply, and one

also can obtain rather reliable error estimates for this procedure. This algorithm has been used for the study of liquid–vapor transitions [6.25, 6.27] and of liquid–liquid phase separation [6.26, 6.97]. For simplicity, we follow [6.9] using the language appropriate for a liquid–vapor transition, and simply take the number of particles N as the order parameter. We are interested in the distribution $P(N)$ for a given choice of variables V, T and μ in the grand-canonical ensemble, at temperatures T distinctly lower than the critical temperature. Then, $P(N)$ has two peaks: one for particle numbers close to $N_v = V\rho_v$ and another close to $N_\ell = V\rho_\ell$, where ρ_v, ρ_ℓ are the densities of coexisting vapor and liquid in the thermodynamics limit, provided μ is close to μ_{coex}, and in between N_v, N_ℓ there occurs a region with a deep minimum (precisely at μ_{coex} we know that $P(N) \propto \exp[-Vf(\rho)/k_{\text{B}}T]$, with $f(\rho)$ shown in Fig. 6.1, for example). However, to sample the distribution efficiently, it is better to sample the reweighted distribution $P_{\text{sim}}(N)$

$$P_{\text{sim}}(N) = P(N)\exp[-w(N)], \tag{6.19}$$

where in the optimum case $w(N) = \ln P(N)$, because then $P_{\text{sim}}(N)$ would be perfectly flat. We note that this is the same idea as used in multicanonical sampling (see Sect. 6.3.) and Wang–Landau sampling (see Sect. 6.4), of course. However, $P(N)$ initially is completely unknown.

The strategy [6.9] now is to divide the region of all particle numbers from $N = 0$ to some $N = N_{\max}$ into m overlapping windows of width ω, and investigate one small window after the other. A histogram $H_K(N)$ monitors how often each state N is visited in the window $[k\omega, (k+1)\omega]$, We denote the values of the kth histogram at its left and right boundary by $H_{k\ell} \equiv H_k(k\omega)$, $H_{kr} \equiv H_k[(k+1)\omega]$, and define the ratio r_k as

$$r_k = H_{kr}/H_{k\ell}. \tag{6.20}$$

After a predetermined number of Monte Carlo steps per window, the (unnormalized) probability distribution is estimated recursively

$$\begin{aligned} P(N)/P(0) &= (H_{0r}/H_{0\ell})(H_{1r}/H_{1\ell})\ldots(H_k(N)/H_{k\ell}) \\ &= \Pi_{i=1}^{k-1} r_i (H_k(N)/H_{k\ell}) \end{aligned} \tag{6.21}$$

with $N \in [k\omega, (k+1)\omega]$. Probability ratios in this equation correspond to free energy differences. Care is needed at the boundaries of a window to fulfill detailed balance: If a move attempts to leave the window, it is rejected, and the previous state is counted once more for the histogram. The number of Monte Carlo steps (count of insertion or deletion trials) is not increased, however, as these moves neither contribute to the statistics.

As one samples one window after the other, a weighted simulation amounts to replace $H_k(N)$ by $H_k(N)\exp[(w_k(N)]$, and the weight function $w_k(N)$ is estimated by extrapolation: After $P(N)$ has been estimated, $w(N)$ is extrapolated quadratically into the next window. The first window is usually unweighted. Virnau and

Müller [6.9] also give evidence (supported by phenomenological arguments) that the statistical errors are basically independent of window size, and carefully discuss the propagation of systematic errors. An application of this technique was already shown in Figs. 6.1 and 6.2.

6.3 Multicanonical Sampling and Other "Extended Ensemble" Methods

Let us consider the problem of thermally driven first-order transitions, as they occur for instance in the case of the well-known q-state Potts model [6.98]. As discussed already in Chap. 2, the probability distribution $P_L(E)$ of the internal energy in a simulation using lattices of size $L \times L (d = 2)$ or $L \times L \times L$ ($d = 3$ dimensions) in the transition region has two rather sharp peaks, separated by a deep minimum in between (cf. also the Ansatz equation (2.99)). When the peaks have equal weight [6.99, 6.100] ("equal weight rule"), the transition from the ordered phase (stable for $T < T_c$) to the disordered phase (stable for $T > T_c$) occurs, giving rise to a rather sharp peak in the specific heat, which develops into a delta function (representing the latent heat of the transition) as $L \to \infty$.

It was rather obvious from the data presented in Chap. 2 that it is very difficult to obtain accurate date on this specific heat peak by naive Monte Carlo sampling, because for large L the transitions between the two peaks of $P(E)$ occur very rarely, but it is necessary to sample many such transitions to estimate the relative weights for these peaks. In Chap. 4, it was described that by single histogram reweighting [6.101] and multiple histogram reweighting [6.102] the accuracy can be significantly improved, since Monte Carlo data yielding a distribution $P_L(E)$ for one inverse temperature β can be reweighted to neighboring temperatures in an interval $\Delta\beta \propto 1/\sqrt{N}$ (where $N = L^d$ here). Nevertheless, if the sampled distribution has a double-peak structure, the problem of estimating correctly the relative weights of the peaks remains.

The multicanonical Monte Carlo method [6.49–6.54] addresses this problem noting from (6.10) that

$$P_L(E) = g_L(E) \exp(-\beta E) \tag{6.22}$$

and suggesting to choose a reweighting with the weight function

$$w_{\text{muca}}(E) \propto 1/g_L(E) \quad \text{for} \quad E_{\min} \le E \le E_{\max}, \tag{6.23}$$

where the interval $[E_{\min}, E_{\max}]$ is the range of energies of interest for the considered transition. Of course, if $g_L(E)$ were known, the problem would be solved, the specific heat could be calculated from $g_L(E)$ analytically, and no Monte Carlo sampling would be required. But $g_L(E)$ is not known, one can only try to find a "working approximation" [6.103] of the weight function $w_{\text{muca}}(E)$. After such a function has

been constructed, one performs an actual Monte Carlo simulation with this function, to obtain a final estimate of $g_L(E)$. Then, one can use (6.22) to obtain averages at the desired values β of interest.

The drawback of multicanonical Monte Carlo is that finding a suitable approximate weight function is not straightforward. (Note that the problem is essentially the same as in standard umbrella sampling: This is no surprise at all, since multicanonical Monte Carlo can be viewed as a reformulation of umbrella sampling [6.64, 6.95, 6.103].) In the case of the Potts model, it turned out useful to sample $P_L(E)$ for a small lattice (where the minimum in between the coexisting phases is not yet very deep), and obtain $g_L(E)$ accurately for small L, using then finite size scaling concepts (cf. Chap. 2) to predict an approximation for $g_L(E)$ at a large value of L. After having then found a good estimate for the actual $g_L(E)$ at this value of L, one can go on to still larger lattices. We refer the reader to the original papers for details [6.49–6.54]. This method can also be generalized to other variables (e.g., studying the first-order phase transition of an Ising model varying the magnetic field [6.14], or varying the chemical potential difference in the case of phase-separating polymer solutions [6.12, 6.24]).

For systems with rugged free energy landscapes, finite size scaling methods for estimating weight factors $w_{\mathrm{muca}}(E)$ often do not work; for example, for biomolecules one particular size is only of interest, or for spin glasses, the weights change too much when the system size is increased. If one does not want to rely on "ad hoc per hand estimates" [6.103], one must try to use a systematic recursive method [6.103–6.105]. Writing the weight of state α as

$$w(\alpha) = \exp[-b(E_\alpha)E_\alpha + a(E_\alpha)] \equiv [-S(E_\alpha)], \tag{6.24}$$

where S can be interpreted as the associated entropy and assuming that the energy spectrum is discrete (step size ε), one can derive a recursion [6.54]

$$b_{n+1}(E) = b_n(E) + \hat{g}_n(E)[\ln H_n(E+\varepsilon) - \ln H_n(E)]/\varepsilon, \tag{6.25}$$

$$\hat{g}_n(E) = g_n(E)/[h_n(E) + g_n(E)], \tag{6.26}$$

$$g_n(E) = H_n(E+\varepsilon)H_n(E)/[H_n(E+\varepsilon) + H_n(E)], \tag{6.27}$$

$$h_{n+1}(E) = h_n(E) + g_n(E), \quad h_0(E) = 0, \tag{6.28}$$

$H_n(E)$ being the histogram entry at the nth step of the recursion. See the monograph by Berg [6.54] for a justification of (6.25)–(6.28). We emphasize, however, that a much more straightforward construction of weights for a multicanonical simulation is achieved by the use of Wang–Landau sampling [6.56–6.68], see next section.

We mention at the end of this section a much more straightforward sampling strategy to create an extended ensemble, which is known as "parallel tempering" or "replica exchange Monte Carlo" [6.65, 6.66, 6.68]. One performs a set of m

canonical Monte Carlo simulations at different values β_i, where $i = 0, \ldots, m-1$, with $\beta_0 < \beta_1 < \cdots < \beta_{m-2} < \beta_{m-1}$.

For these runs standard Boltzmann weight factors $\exp[-\beta_i \mathcal{H}(\vec{x}]$ are used. From time to time, one attempts to exchange neighboring β values,

$$\beta_i \to \beta_{i-1}, \quad \beta_{i-1} \to \beta_i, \quad i = 1, \ldots, n-1 \tag{6.29}$$

at fixed configuration of the systems, just treating the exchange in (6.29) as an additional type of Monte Carlo move, which is subjected to the standard importance sampling acceptance criterion. Of course, the set of values $\{\beta_i\}$ has to be chosen such that a reasonably large acceptance rate results. Obviously, this requires that the distributions $P_L(E)$ for β_i and β_{i-1} have a strong overlap. Thus, the number m of neighboring temperatures must scale like $m \propto \sqrt{N}$ as $N \to \infty$, but nevertheless the method is useful, particularly for systems such as glasses and spin glasses [6.29, 6.31, 6.41–6.46].

6.4 Wang–Landau Sampling

The Wang–Landau algorithm [6.56–6.63] is an elegant iteration method to construct directly the energy density of states, $g(E)$, in (6.10). This method has the merit of great simplicity and therefore finds widespread applications [6.61]. The idea is to perform a random walk in energy space with a flat histogram. One can also generalize this method to other variables (e.g., sample for an Ising ferromagnet $g(E, M)$ where M is the magnetization), but such generalizations will not be considered here. Unlike histogram methods where one extracts estimates of $g(E)$ from probability distributions generated by standard Monte Carlo simulations, temperature plays no role in this algorithm at all! Transitions between microstates (e.g., spin flips in an Ising model) are carried out according to probability

$$p(E \to E') = \min\{g(E)/g(E'), 1\} \tag{6.30}$$

If $g(E)$ where known in beforehand, a Markov process based on (6.30) would generate a flat histogram. Since $g(E)$ is not known in before hand, an iteration process needs to be performed to construct $g(E)$, starting from a simple initial guess. In the absence of any a priori knowledge on $g(E)$, it is natural to choose as an initial density of states simply $g(E) = 1$, for all E. For simplicity, we restrict here attention to a case where the energy spectrum is discrete and bounded ($E_{\min} \le E \le E_{\max}$), as it is the case for an Ising model or a lattice model for a polymer chain, but we emphasize at the outset that such a restriction is not necessary, and many successful applications of Wang–Landau sampling to off-lattice models (like fluids or polymer chains with Lennard–Jones interactions) have been given [6.61].

Since the actual $g(E)$ is very different from $g(E) \equiv 1$, it is necessary to have a procedure that leads quickly away from this initial condition. This is achieved by

replacing $g(E)$ by $g(E)$ times f, where the initial guess of the modification factor f_1 is $f_1 = e^1$, whenever a state with energy E is visited. During this random walk in energy space controlled by (6.30), one accumulates a histogram $H(E)$; that is, starting out with $H(E) = 0$ for all E, one replaces $H(E)$ by $H(E) + 1$ whenever E is visited. The moves that are carried out to realize (6.30) are just the standard moves that one also would use in a standard Monte Carlo simulation that attempts to sample the Boltzmann distribution (e.g., spin flips of an Ising model, random displacements of an effective monomer to a new position in a model for a polymer chain). This sampling process of the histogram is continued, until the histogram $H(E)$ is "reasonably flat". In practice, it has turned out that a useful criterion for this "flatness" is to require that the minimum entry in the histogram is not smaller than 80% of its mean value [6.56–6.58].

When this flatness has been achieved in the first iteration step, one resets all histogram entries $H(E)$ to zero and performs a sampling with a modification factor $f_2 = \sqrt{f_1}$ (and in the ith step, one chooses $f_i = \sqrt{f_{i-1}}$). Otherwise, the procedure is identical to the procedure followed in the first step. This procedure is iterated until f_i reaches a minimum value, which in practice is chosen to be of order $f_{\min} \approx \exp(10^{-8})$. While the detailed balance principle clearly is violated in the early stages of the iteration, this is no problem in the final run.

The final density of states $g(E)$ then is used to calculate averages, for example, the specific heat

$$C(T) = (\langle E^2 \rangle_T - \langle E \rangle_T^2)/(k_B T^2), \tag{6.31}$$

where moments $\langle E^k \rangle$ are straightforwardly computed from $g(E)$ as

$$\langle E^k \rangle = \sum_E E^k g(E) \exp(-\beta E) / \sum_E g(E) \exp(-\beta E). \tag{6.32}$$

Of course, it is possible to use the run with the final choice $f_{\min}$ to obtain more detailed information. For example, in the study of Ising models, one would like to estimate moments of the magnetization distribution $\langle m^k \rangle_T$. Sampling of the full (two-dimensional) joint density of states $g(E, m)$ would solve this problem, but for large systems this would require a huge computational effort. Typically, it is sufficient (and much easier) to only take "microcanonical" averages, $\langle |m|^k \rangle_E$, from which then the desired canonical averages follow,

$$\langle |m|^k \rangle_T = \sum_E \langle |m|^k \rangle_E \; g(E) \exp(-\beta E) / \sum_E g(E) \exp(-\beta E), \; k = 1, 2, 3, \ldots . \tag{6.33}$$

As an example, we recall a recent study of first-order interface localization–delocalization transitions in thin Ising films with competing boundary fields [6.106]. Here, one considers the Ising Hamiltonian on a simple cubic lattice in a $L \times L \times D$ geometry, with periodic boundary conditions in x and y directions only. In the z-direction, one has a free boundary condition, and in addition surface magnetic fields H_1 and $H_D = -H_1$ act on the first and last layer of the film, in the z-direction. In these surface planes $n = 1$ and $n = D$, one also chooses exchange interactions J_s

different from the exchange J in the bulk. Thus, the Hamiltonian is ($S_i = \pm 1$)

$$\mathcal{H} = -J \sum_{\langle i,j \rangle \text{bulk}} S_i S_j - J_s \sum_{\langle i,j \rangle \text{surfaces}} S_i S_j - H_1 \sum_{i \in n=1} S_i - H_D \sum_{i \in n=D} S_i \,. \quad (6.34)$$

Studying this system for temperatures below the critical temperature T_{cb} of the bulk Ising models, there occurs a regime of temperatures $T_c(D) < T < T_{cb}$ where the total magnetization of the thin film is zero, because two domains of equal size but opposite magnetization occur, stabilized by the surface magnetic fields of the same sign and separated by an interface in the center of the film. However, for $T < T_c(D)$, the interface gets bound to either the left or the right wall, and hence one has a nonzero magnetization when the interface is localized near one of the walls, since then one of the domains is much larger than the other. When J_s/J exceeds a critical value, the transition is of first order, and then a study by finite size scaling methods is very difficult [6.107].

Schulz et al. [6.106] applied the Wang–Landau algorithm for the choice $H_1/J = 0.25$ and $J_s/J = 1.5$, using thicknesses $D = 6, 8$ and 12 lattice spacings, while the lateral linear dimension was varied up to $L = 128$. Figure 6.5 shows typical data for the fourth-order cumulant $U_4 = 1 - \langle m^4 \rangle / 3 \langle m^2 \rangle^2$ plotted versus inverse temperature. At a first-order transition, this quantity is predicted [6.108] to diverge towards minus infinity proportional to $-L^2$ at the first-order transition, and indeed a deep and very sharp minimum develops (Fig. 6.5). Unfortunately, it was found that already for $D = 12$ (and $L = 128$) the convergence of the Wang–Landau algorithm turned out to be problematic, due to "entropic barriers" in the phase space [6.106].

Such entropic barriers were first pointed out and carefully analyzed by Neuhaus and Hager [6.109] for the case of the two-dimensional Ising model. They showed that both the multicanonical Monte Carlo method and the Wang–Landau Algorithm suffer from this problem in a similar way; if entropic barriers are too high, the configuration

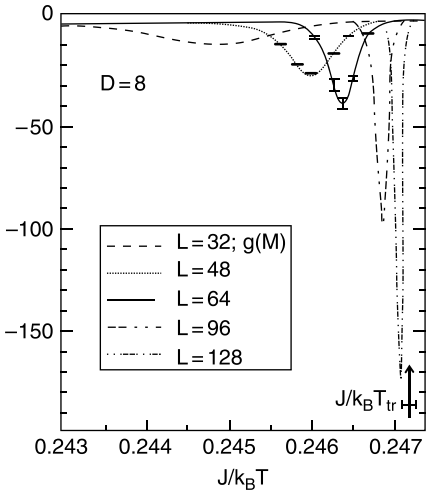


Fig. 6.5 Reduced fourth-order cumulant U_4 of a thin Ising film of thickness $D = 8$ plotted versus inverse temperature for different linear dimensions L (note that for $L = 32$ a full histogram $g(M)$ of the magnetization was employed). For $L = 48$ and $L = 64$ multiple iterations were performed, allowing to obtain error estimates, as indicated. The *arrow* indicates the estimate of the extrapolated critical temperature, $J/k_B T_c(D)$

space of the system is insufficiently sampled, and this lack of ergodicity leads to results that are systematically wrong. In the case of the $d = 2$ Ising model transition from positive to negative magnetization, the entropic barriers occur in the two-phase coexistence region, for example, when a transition from a droplet-type configuration to a slab-like domain configuration should take place (analogous to the transition discussed in Figs. 6.1 and 6.2). The "transition state" that amounts to the barrier has a lens-type shape of a very elongated droplet (with a linear dimension L in its long directions) [6.109], and it is very unlikely to reach such a state with the standard Monte Carlo moves, which do not involve any bias towards such a state.

Entropic barriers have the effect that the time to sample the full equilibrium density of states grows exponentially with the linear dimension L of the system [6.109]. For both the multicanonical and the Wang–Landau algorithm even without entropic barriers, there is already a large effort required, since the energy space to be sampled increases proportional to the particle number, $E_{\max} - E_{\min} \propto N$, and hence performing a simple random walk type motion one would predict that the relaxation time scales like N^2. Practical experience shows, however, that some correlation effects do occur in this diffusion process, so the actual relaxation time increases even faster [6.60, 6.74]. In practice, it is advantageous to divide the energy interval $E_{\max} - E_{\min}$ into many subintervals, in which Wang–Landau sampling can be carried out in parallel (but care must be taken to deal with the boundaries of these intervals correctly [6.59]).

Thus, the judgement of accuracy for the Wang–Landau algorithm (and similarly, for the multicanonical method) is somewhat subtle. It is always advantageous to carry out multiple independent runs and try to estimate the errors from these. However, there are cases where lack of ergodicity really is a problem, such as problems involving very dense configurations of polymer chains (see, e.g., Fig. 6.3). Good results are only obtained if the set of Monte Carlo moves is large enough so the system does not get trapped in such configurations.

In view of all these problems, it is not surprising that in spite of these extended sampling methods problems such as spin glasses, structural glass transition, protein folding, are still heavily debated, but interesting progress clearly has been obtained [6.29].

6.5 Transition Path Sampling

Here we return to the problem already alluded to in Fig. 6.3. Suppose the system is stable or metastable in the basin near the starting points (white dots) of the two trajectories, that is, typically it will stay there for a very long time. Only very rarely it will follow a path over high saddle points in a complex free energy landscape to make a transition into another deep basin (arrows).

The aim of transition path sampling then is to gain information on all probable path ways, in order to analyze the transition mechanism. For example, in the context of chemical reactions, one would like to identify the "reaction coordinate" [6.77];

in the case of crystal nucleation from the melt, one wants to get information on the crystal, structure size, and shape of the nucleus that forms [6.110]; etc. This "reaction coordinate" need not be a scalar, but may contain several variables: For example, in nucleation from an unmixing binary fluid mixture one expects that both the size of the droplet and its composition change during the growth of the droplet from subcritical to supercritical.

Since the pathways collected with transition path sampling are trajectories consistent with the markovian master equation description of Monte Carlo sampling, the full information on the kinetics of the transition can be extracted [6.83]. In fact, this technique and its ramifications [6.84–6.90] has taken an impressive development since its invention, and we can give here only the flavor of the approach, and refer to the quoted literature for details of its implementation. We also emphasize that transition path sampling can be adapted to cases where the underlying dynamic simulation method is not Monte Carlo but Molecular Dynamics or Brownian Dynamics [6.83]. However, also in these cases the weight of the trajectories in the "transition path ensemble" is obtained by Monte Carlo methods. The idea is to carry out a random walk in the space of trajectories rather than in configuration space. The basic step generates a new path $\{\vec{x}^{(n)}(\{t\})\}$ from an old one $\{\vec{x}^{(0)}(\{t\})\}$. The underlying dynamics of the system that is simulated defines a "path ensemble" $P_{\mathrm{AB}}\{\vec{x}(\{(t)\}\}$, where we denote the initial state as A and the final state as B. The initial condition is prepared by placing the system in a heat bath at temperature T, and hence the distribution of initial conditions is just the standard canonical distribution, but some variable is constrained so that the state point falls in the region of one minimum in the free energy landscape (Fig. 6.3) where all trajectories start.

In order that all trajectories are compatible with the path ensemble, one introduces a transition probability $p[\{\vec{x}^{(0)}(t)\} \to \{\vec{x}^{(n)}(t)\}]$ that satisfies the detailed balance condition with $P_{\mathrm{AB}}\{\vec{x}(\{(t)\})\}$,

$$P_{\mathrm{AB}}\left\{\vec{x}^{(0)}(\{t\})\right\} p\left[\left\{\vec{x}^{(0)}(t)\right\} \to \left\{\vec{x}^{(n)}(t)\right\}\right] = P_{\mathrm{AB}}\left\{\vec{x}^{(n)}(\{t\})\right\} p\left[\left\{\vec{x}^{(n)}(t)\right\} \to \left\{\vec{x}^{(0)}(t)\right\}\right]. \tag{6.35}$$

The transition between individual states (at one time t of the Markov process, $\vec{x}^{(0)}(t) \to \vec{x}^{(n)}(t)$) hence is replaced by a transition between two full trajectories, such as shown in Fig. 6.3.

This probability p then can again be written as a product of a "generation probability" of a new path and an "acceptance probability", and from (6.35) one then readily can postulate a Metropolis importance sampling rule. Of course, the subtle problem is the generation of new paths with a reasonably high acceptance probability. To solve this problem, one mostly relies on the so-called "shooting algorithm" [6.28]. For this purpose, the path is divided into many small time slices. From the randomly selected time slice t', one carries out a move to a new state, $\vec{x}(t') \to \vec{x}'(t')$ according to the rules of the underlying dynamics. From the new state $\vec{x}'(t')$, forward trajectories (that end up in B) and backward trajectories (that end up in A) are generated, so that one can glue one forward and one backward trajectory together to get a new full trajectory going from A to B. This new trajectory eventually is

accepted (or rejected), depending on the Metropolis test based on (6.35). Of course, in order to be able to start such a sampling, an initial trajectory must be available. Just as in ordinary Monte Carlo sampling, where the initial state does not need to be similar to the states characteristic for the final thermal equilibrium, for example, one can start the simulation of an Ising ferromagnet always from a perfectly ordered spin configuration irrespective of the temperature of the simulation, one can start transition path sampling from a completely atypical trajectory, and then try to relax the system to converge towards the dominating trajectories. Of course, it depends on the model chosen whether such an approach is practically feasible or not (and also the task of finding an initial trajectory is a nontrivial task!).

Before proceeding, we discuss a simple example, nucleation in the three-dimensional Ising model with nearest-neighbor exchange [6.94]. Pan and Chandler [6.94] chose a temperature $T = 0.6T_c$, a lattice of linear dimension $L = 32$, and the magnetic field chosen $H = -0.55$, but the initial state was chosen to have positive magnetization, constraining the system such that the maximum size of the clusters of overturned spins was $N_A = 26$. Since classical nucleation theory (using the known interfacial free energy of the Ising model) yields critical cluster size $N^* = 200$, the "reactant" state which should contain a supercritical cluster was chosen to have $N_B = 260$ overturned spins in the largest cluster (the actual critical cluster size was found to be near $N^* = 115$, so N_B was chosen large enough). An initial trajectory was obtained by generating large overturned clusters with umbrella sampling, and running Monte Carlo trajectories from such states, until a path connecting A to B could be built. Even if this initial path were uncharacteristic for the transition path ensemble, equilibration was no problem, since this trajectory was relaxed with 25,000 moves before the sampling was started, and then 1000 independent trajectories were generated (one every 100 moves). From these trajectories, one can also construct the so-called "transition state ensemble" [6.28]: This is the ensemble of states $\{\vec{x}\}$ from which 50% of the trajectories lead back to A, while 50% lead forward to B. Figure 6.6 shows histograms of the cluster size N^* and surface area S^* in this transition state ensemble [6.94], and hence illustrates that the concept of a well-defined size of a critical droplet (made by classical nucleation theory [6.111]) clearly is a severe simplification, as expected [6.93].

A useful concept in transition path sampling is the "committor distribution". It is defined as the fraction of trajectories started in A and going via a state $\vec{r}$ to reach the state B after a time t [6.83]

$$P_B(\vec{r}, t) \equiv \frac{\int \mathcal{D}\vec{x}(t') P(\{\vec{x}(t')\}) \delta(\vec{r}_0 - \vec{r}) h_B(\vec{x}(t))}{\int \mathcal{D}\vec{x}(t') P(\{\vec{x}(t')\}) \delta(\vec{r}_0 - \vec{r})}. \tag{6.36}$$

Here, $h_B(\vec{x}) = 1$ if $x \in B$ and zero else, paths that start at time $t = 0$ at $\vec{r} = \vec{r}_0$ but do not end up at time t in the region B are not included in the numerator, while they are included in the denominator. Thus, the committor distribution is a statistical measure for "how committed is a given configuration to reach the 'product state' B" [6.83]. In an practical simulation, only a finite sample of $\mathcal{N}$ trajectories started at $t = 0$ at a point $\vec{r}$ in configuration space is available, and then (6.36) can be written simply as

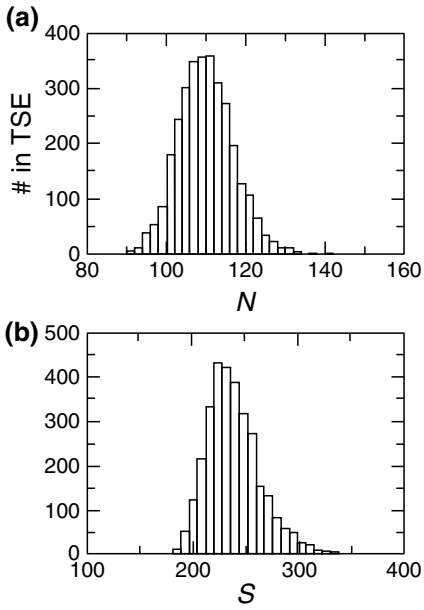


Fig. 6.6 Distribution of cluster size (a) and surface areas (b) in the transition state ensemble, obtained from transition path sampling [6.94], for a nearest neighbor simple cubic Ising ferromagnet at 60% of its critical temperature and a field $H/J = -0.55$

$$p_{\mathrm{B}}(\vec{r}, t) \approx \frac{1}{\mathcal{N}} \sum_{i=1}^{\mathcal{N}} h_{\mathrm{B}}(\vec{x}_i(t)). \tag{6.37}$$

Analogously one can define $p_{\mathrm{A}}(\vec{r}, t)$, and a formal definition of the "transition state ensemble" then becomes [6.28, 6.77–6.83]

$$p_{\mathrm{A}}(\vec{r}) = p_{\mathrm{B}}(\vec{r}), \tag{6.38}$$

where we also have invoked a time-independent form of the committor (counting all events that first reach A or first reach B, irrespective how long it has taken).

While transition path sampling is rather straightforward to implement in the example of nucleation in the kinetic Ising model (Fig. 6.6), where one has a qualitative insight into the kinetics of the process and the nature of transition states a priori, there are problems where no such knowledge is available, for example, in the folding process of complex off-lattice models for proteins [6.34]. In such cases, it is useful to combine then the transition path sampling with complementary sampling strategies to obtain information on the free energy landscape, for example, the replica exchange method, that has been mentioned in Sect. 6.3 [6.34]. Since at the time of writing "for large proteins the computational effort due to both system size and long time scales becomes prohibitive" [6.34], this subject will remain an area of further research.

6.6 Concluding Remarks

In this chapter, methodic advances in Monte Carlo sampling have been described that overcome, to some extent, the limitation of the standard importance sampling algorithm that no information on the partition function and hence the free energy is available. This limitation has hampered the study of cases where several thermodynamic states of the system (corresponding to valleys or basins of attraction in the free energy landscape of the system) compete with each other and are separated by large barriers. The standard example of this situation are first-order phase transitions, but other cases where the free energy landscape is "rugged" and the order parameters distinguishing the various basins is not well understood (spin glasses, proteins, etc.) may be even more interesting. The methods that yield information on the relative weights of these states described by the basins often yield also information on the height and nature of the free energy barriers between them, a question which is interesting in its own right. In fact, transition path sampling then addresses also the question of the kinetics of the pathways along which such barriers are crossed.

Many of the techniques described here start out from the old [6.112] and well-known (e.g., Fig. 8 of [6.91]) fact that Monte Carlo sampling gives not only information on averages such as written in (6.1) and (6.3), but also *distribution functions* are obtained which contain valuable information. While the use of distribution functions to justify the finite scaling analysis of first- and second-order phase transitions (see Chaps. 2–4 of this book) now is standard practice, including histogram reweighting methods [6.101, 6.102], only since about 1991 [6.49] the problem of free energy barriers (that may contain, among other things, information on interfacial free energies [6.21]) has become the focus of methodic development of sampling strategies. The old idea of umbrella sampling [6.47] has been rediscovered, and further developed, in many different variants (e.g., [6.9, 6.49–6.55, 6.66–6.73], and much interesting insight into various problems of statistical thermodynamics of condensed matter has been gained. A particularly simple but efficient approach to sample the energy density of states, Wang–Landau sampling, has been proposed 10 years later, in 2001 [6.56, 6.57], nevertheless is already the most widely used method in many different contexts, from the accurate estimation of complex phase diagrams of magnetic systems [6.113] to membrane proteins [6.114]. In view of the rapid development of methods during the last two decades, it is likely that further improvements of methodology will occur in the near future, but now the methods described here should allow a wealth of further useful applications.

Of course, other methodic aspects outside the scope of this chapter (e.g., cluster algorithms for spin models with long range interaction [6.115] and for off-lattice fluids [6.116], finite size scaling for asymmetric fluid criticality [6.117] and – last but not least – Quantum Monte Carlo methods, which even have started to compete with quantum chemistry methods [6.118–6.121]) have also seen major progress. Thus, Monte Carlo simulation in statistical physics continues to gain importance.

Chapter 7
Rejection-Free Monte Carlo

7.1 Introduction

So far, we have been using the rejection Monte Carlo algorithms. To remind us, the algorithms proceed from state x to possible state x' as outlined in Algorithm 1.

Algorithm 1 Accept/Reject Monte Carlo Algorithm

1: Choose initial state x
2: **for** n-of-samples **do**
3: Select a new state x'
4: With probability p accept, i.e. set $x = x'$
5: With probability $(1 - p)$, x' is rejected
6: **end for**

The probability will depend on some change induced by the state change as for example in the case of the Metropolis-Hastings Monte Carlo. If we are to construct Monte Carlo methods that do not involve any form of accept/reject criterion in the sense outlined above and as it was used in the previous chapters, then we have to select states (or events) that for sure will be accepted. Thus, the methods will be (synchronously or asynchronously) event-driven [7.1] (see Algorithm 2).

Algorithm 2 Event-Driven Algorithm

1: **for** n-of-samples **do**
2: Identify all possible events
3: Identify the event with the smallest time stamp Δt
4: Set time $t = t + \Delta t$
5: **end for**

Here we will expose methods that rely on rates between states thus the sequence that ultimately will be generated evolves in time (see Fig. 7.1). However, not as in the

K. Binder and D. W. Heermann, *Monte Carlo Simulation in Statistical Physics*,
Graduate Texts in Physics, https://doi.org/10.1007/978-3-030-10758-1_7

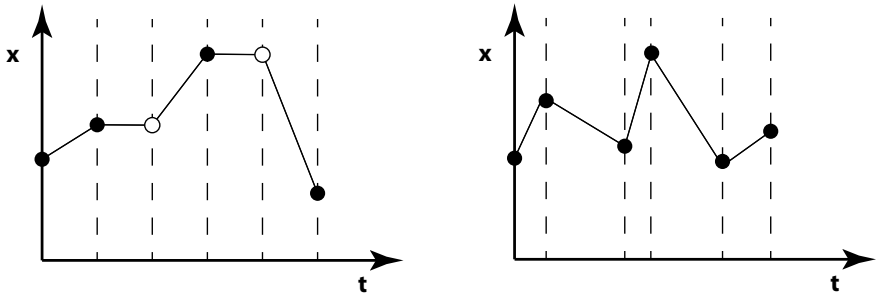


Fig. 7.1 The lhs panel shows the typical systolic propagation of time for example in the Metropolis Monte Carlo. Sometimes new state proposals are rejected (circles) and the previous state is the new state. The rhs panel depicts the leaps in time that are made to achieve a rejection free algorithm

previous chapters systolically, driven by a constant increment in time, but by leaps of various length in time. This also opens up the possibility to make rigorous the notion of time in Monte Carlo methods.

7.2 Rejection-Free Methods

Consider a state space Ω and a sequence $\{x_{t_k} \in \Omega\}$ of states from the state space. Often we simply write i or j etc. to label the states. Here we assume $t_0 < t_1 < \cdots < t_k < \cdots$. So far we have had $\Delta t = t_k - t_{k-1}$ constant, i.e. the system was moved forward in time by a constant stride. Furthermore, for two states (x_{k-1}, x_k) we have the Markov property so that the sequence $\{x_{t_k} \in \Omega\}$ is a Markov chain.

Let us now look at continuous-time Markov chains $\{x_t \in \Omega | t \in \mathbb{R}, t \geq 0\}$. For the chain to be a continuous-time Markov chain the following condition needs to apply

$$\mathbb{P}(x(t+\tau) = j | x(\tau) = i, x(u) = k, 0 \leq u \leq \tau) = \mathbb{P}(x(t+\tau) = j | x(\tau) = i) \; . \tag{7.1}$$

Define

$$p_{ij}(t) := \mathbb{P}(x(t+\tau) = j | x(\tau) = i) = \mathbb{P}(x(t) = j | x(0) = i) \tag{7.2}$$

and for any state i we have (for N possible states)

$$\sum_{j=1}^{N} p_{ij}(t) = 1 \; . \tag{7.3}$$

Let $P(0) = \lim_{t \searrow 0} P(t) = I$ be the initial condition. Then the matrix R defined by

$$\lim_{h \searrow 0} \frac{P(h) - I}{h} = P'(0) = R \tag{7.4}$$

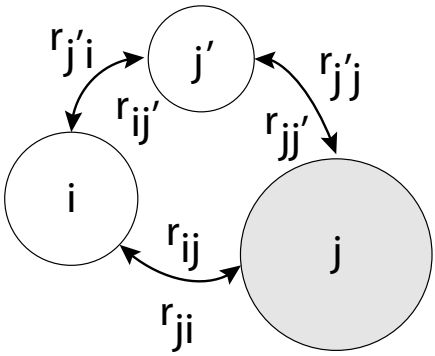


Fig. 7.2 The figure shows the general situation where the circles denote states in state space that belong to the same state i

is the infinitesimal generator of the continuous-time Markov process with rate r_{ij}

$$\sum_{j=1, j\neq i} r_{ij} = -r_{ii} \tag{7.5}$$

and

$$r_{ij} = \lim_{h \searrow 0} \frac{p_{ij}(h)}{h} \geq 0 \text{ and } r_{ii} \leq 0 \, . \tag{7.6}$$

Define $r_i := -r_{ii} > 0$ to be the rate corresponding to state i. Given R, then for all $t \geq 0$

$$P'(t) = RP(t) \, . \tag{7.7}$$

and

$$P(t) = Re^{-Rt} \tag{7.8}$$

as the first-passage-time distribution and further

$$p_{ij} = r_{ij} e^{-r_{ij} t} \, . \tag{7.9}$$

Since we are talking about first-passage only, only one of the possibilities can happen.

Thus, rather than focusing on the transition probabilities (c.f. Fig. 7.2) as we have in the previous chapters, we can focus on the rates between states opening up to models where there is no Hamiltonian. Even more so, the rates themselves may depend on time. If they do not then the Markov process is stationary.

Let n_i denote the population of state i. Given that we are dealing with a thermal system then n_i must be proportional to $\exp\{-F(i)/k_{\mathrm{B}}T\}$. In equilibrium if we have detailed balance then

$$n_i r_{ij} = n_j r_{ji} \, . \tag{7.10}$$

Thus, what is needed for a model is a state space Ω and a set of rates R, i.e. (Ω, R). This can for example be a set of chemical reactions with the corresponding rates.

We envisage that at any given time for a state i not all states j are accessible. Thus it is convenient to relabel the currently accessible states with a new label. We arrive at a list of N possible events and a list with corresponding rates

$$\{E_n \in \Omega\} \quad \text{with } n = 1, \ldots, N \tag{7.11}$$

$$\{r_n\} \quad \text{with } n = 1, \ldots, N \tag{7.12}$$

From a computational point of view, it is immediately apparent that what is needed is a well performing bookkeeping algorithm for the events and the rates as they may change after an event has been chosen (see later).

Let us consider this for the Ising model. It was pointed out by Bortz et al. [7.2] that the probability of accepting new configurations in the Ising model is very low in some cases. Consider the case when the temperature is low. Then two spins will have likely the same orientation and thus a reversal has very low probability. Thus, out of the N attempts only a very low fraction will result in changes. Suppose only attempts are made that are successful. For this the rates r_{ij} from state i to j need to be known a priori.

In the Ising case we know transition rates among states a priori. For the two-dimensional Ising model

$$\mathcal{H} = -J \sum_{<i,j>} S_i S_j \quad S_i = \pm 1 \tag{7.13}$$

with its spin-up spin-down symmetry we have the situations as shown in Table 7.1. Altogether we have ten possible states, depending on the number of neighbors the central spin is surround by. Each of these states we assign a class.

Assume further that the transition probability between states is given by

$$p = \frac{x}{1+x} \quad \text{with} \quad x = \exp\{-\Delta\mathcal{H}/k_\mathrm{B}T\}\,, \tag{7.14}$$

then all possible transitions $r_{i,}$ are given.

One possibility is to use (7.9) to draw time increments for the event to happen. (This algorithm is known as the first-reaction method [7.3]). For this we generate a random number $\rho \in (0, 1)$ and compute

$$t_{ij} = -r_{ij}^{-1} \ln(\rho)\,. \tag{7.15}$$

Table 7.1 Classes for the kinetic Monte Carlo (n-fold way) algorithm proposed by Bortz et al. [7.2] Corresponding to each class i there is a probability p_i

Spin	↑ (+1)					↓ (−1)				
Neighbours	4	3	2	1	0	0	1	2	3	4
class	1	2	3	4	5	6	7	8	9	10

Thus, for every state change we know the probability and the first passage times. What remains to do is to identify the state change $i \to j$. For this we select the reaction coordinate that comes first in time

$$\Delta t = \min_{ij} t_{ij} \ . \tag{7.16}$$

Then this state change is performed and time advances (see Algorithm 3)

$$t = t + \Delta t \ . \tag{7.17}$$

Algorithm 3 First Reaction Monte Carlo Algorithm

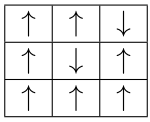
```
Initial time t = 0
Choose initial state i
for n-of-samples do
    Set up list of transition rates r_ij (size N)
    Generate N random numbers ρ_j from a uniform distribution on (0, 1]
    t_ij = r_ij^{-1} ln(ρ_j^{-1})
    Δt = min_ij t_ij
    Carry out event i → j that is minimum
    Update t = t + Δt
    i ← j
end for
```

Hence, we only perform those state changes that actually occur. This is in contrast to the procedure that we have developed in the previous chapters. Note that this algorithm uses $\mathcal{O}(N)$ to build the list of transition rates, $\mathcal{O}(N)$ for the number of random numbers and $\mathcal{O}(N)$ to determine the minimum time.

A further development on this idea was put forward by Gibson and Bruck [7.4] with their next reaction method. The method only uses one random number and time proportional to the logarithm of the number of reactions due to the use of a dependency graph, eliminating costly recalculations of the transition rate list (see also the composition and rejection stochastic simulation algorithm [7.5]).

Another aspect is, that there may be relatively less significant events, i.e., minimal changes in transition rates. This idea has led to the development of the τ-leap algorithm. τ-leaping [7.6, 7.7] requires the knowledge of the rate change. Further, the assumption is made that during the of τ-leap the propensity function is assumed to stay constant during each leap.

The obvious difference to the Metropolis Monte Carlo algorithm is that time does not advance in fixed increments but rather leaps in non-constant strides. It must further be pointed out that the transition probabilities change at every step. Indeed, one of the key features is that the distribution of rates is coupled to the state space [7.8] and can change. For the Ising case there is no such problem. This can be seen when we consider the two-dimensional case shown in Diagram 7.2.



This translates into the class scheme from Table 7.1.

2	3	10
2	4	3
1	2	2

A spin flip can change the transition probability and with it the class.

The origin of the algorithm that has partially been exposed above can be traced back to Young and Elcock [7.9], Bortz et al. [7.2] (for the Ising model, n-fold-way), Gillespie [7.3, 7.10] (chemical reaction) and Cox [7.11]. In their formulation the algorithm, know now as kinetic Monte Carlo the choice of the transition that is going to take place is made slightly different (for reviews see for example [7.12–7.16]). Also the Optimized Direct Method [7.17] and the Sorting Direct Method [7.18] have been developed as variations of the basic idea.

The starting point is a choice of a state the system is started in. This determines the possible states that the system can transition into and the corresponding rates r_{ij}. The next step is to compute the sum over all the possible rates from i to j, i.e. all possible reaction paths. The next step then is to pick one of the possible reaction paths with equal probability followed by advancing the time as shown in Algorithm 4.

Algorithm 4 Kinetic Monte Carlo Algorithm

1: Initial time $t = 0$
2: Choose initial state i at random
3: **for** n-of-samples **do**
4: Set up list of transition rates r_{ij} (size N)
5: Compute $R_{i,j} = \sum_{k=1}^{j} r_{ik}$ for $j = 1, ..., N$
6: Compute $R_i = R_{i,N}$
7: Generate ρ from a uniform distribution on (0, 1]
8: Choose i such that $R_{i,j-1} < \rho R_i \leq R_{ij}$
9: Carry out event j
10: Update $i \rightarrow j$
11: Generate ρ from a uniform distribution on (0, 1]
12: $\Delta t = R_i^{-1} \ln(\rho^{-1})$
13: $t = t + \Delta t$
14: **end for**

The beauty of the kinetic Monte Carlo Method is that it easily generalizes to arbitrary states and reactions. This is why it has been used for many condensed matter systems [7.9, 7.19–7.21] with certain refinements [7.4, 7.18, 7.22–7.24] and coupled

to molecular dynamics [7.25]. Further developments are the coarse-grained kinetic Monte Carlo [7.13, 7.26] and the first-passage kinetic Monte Carlo algorithm [7.27].

Let us return to the initial example of the Ising Model. Let n_i be the number of spins in class i (see Table 7.1), then we need to choose the relative weights $n_i p_i$ according to Algorithm 4 and once a class has been chosen a spin in that class is chosen with probability $1/n_i$.

Fichthorn and Weinberg [7.28] showed that under the condition of detailed balance and the effective independence of the events, the Algorithm 4 yields a Poisson process and that static and dynamic properties are consistent with the Hamiltonian dynamics [7.29]. However, detailed balance is not necessary! As we will see later, the kinetic Monte Carlo method is used for non-equilibrium situation and where detailed balance is not fulfilled but global balance is achieved.

Note that number of operation, i.e. the complexity is $O(N)$. Makysm [7.30] showed that using a binning method and recursive search trees, the complexity can be brought down to $O(\log_2 N)$ [7.23].

For completeness, even though we are in the chapter on rejection-free Monte Carlo, here is a rejection algorithm for the model pair (Ω, Q).

Algorithm 5 Rejection Kinetic Monte Carlo Algorithm

```
for n-of-samples do
    Set up list of transition rates r_n (size N)
    Compute an estimator for the sum of rates r̄
    while state not selected do
        Generate ρ from a uniform distribution on [0, N)
        Compute n = (Int)(ρ) + 1
        Select n if n − ρ < r_n/r̄
    end while
    n is new state
end for
```

7.3 Parallelization

For the Ising model, Lubachevsky [7.31] has succeeded to parallelize the Monte Carlo algorithm based on the ideas put forward in the more general context by Chandy and Misra [7.32, 7.33]. He formulated the algorithm as a distributed discrete-event system. Various methods have been designed specifically with lattice models at focus [2.34–2.36]. Also the scaling properties of these type of algorithms have been investigated [7.37, 7.38] associating the development of the individual time increments at the individual processors with time increments corresponding to depositions and thus identifying this with surface growth (Kardar–Parisi–Zhang equation [7.39]). The parallelization of the τ-leap has been done by Xu et al. [7.40] and for the presence

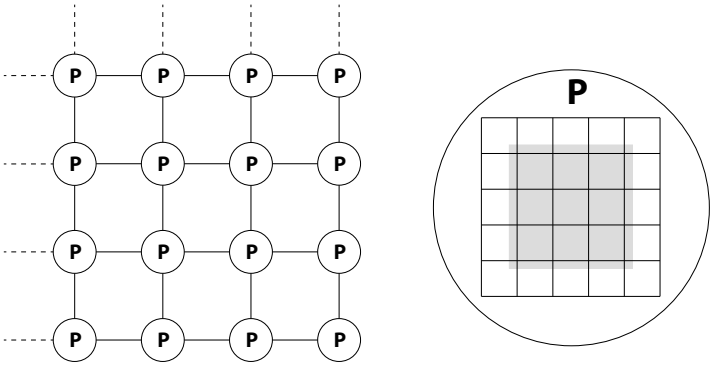


Fig. 7.3 For the simplest parallel kinetic Monte Carlo algorithm, we assume that the topology for the processors is that of a lattice (for simplicity here a simple square lattice with the processors (P) at the nodes of the lattice) with possible periodic boundary condition (dashed lines). The solid lines represent bi-directional communication channels (lhs). The rhs panel shows the possibility that a processor has been assigned more than site, say for the 2-D Ising model, L/l lattice sites. The gray shaded area is the part where no communication between the processors is needed for a decision to flip a spin

of long-range interactions see [7.41]. Also much effort has gone into parallelization of the Gillespie ansatz, for example, Komarov [7.42].

The key problem in the parallelization is to avoid event time incompatibilities with communications. The solution that Lubachevsky [7.31] has put forward is the strict synchronization (c.f. Fig. 7.3 and Algorithm 6). The problem is solved in this algorithm using a global synchronization at the slight expense of efficiency. The algorithm presented here is aiming at the above outlined Ising situation. We assume two functions nextState$(i, t_i, \text{neighbours}(i))$ which calls upon the neighbor processors for the corresponding states s_j and nextTime(t_i) delivers the next time.

Algorithm 6 Lubachevsky Parallel Monte Carlo Algorithm

$s' = s_i$
$t' = t_i$
for n-of-samples **do**
  if $t_i \leq \min_{j \in \text{neighbours}(i)} t_{ij}$ **then**
    $s' = \text{nextState}(i, t_i, \text{neighbours}(i))$
    $t' = \text{nextTime}(t_i)$
    Global synchronize
    $t = t'$
    $s' = s$
    Global synchronize
  else
    Global synchronize
    Global synchronize
  end if
end for

Assume that in the Ising case the lattice is much larger than the number of processors and that there are L/l lattice sites per processors (i.e. $L \times L$ lattice with $l \times l$ blocks). There are now interior and boundary sites to be handled by the Algorithm 6. Korniss et al. [7.43] have argued that the synchronization steps in the algorithm are not necessary. If the same random number generator runs on each of the processors with the same initial seed, they argue that the probability of equal-time nearest-neighbor updates is of measure zero. Thus they suggest to treat the interior spins (gray shaded region in Fig. 7.3) like regular spins and use

$$p = \min\{1, \exp(-\Delta H/kT)\} \tag{7.18}$$

with $\Delta t = -\ln(\rho)$ (ρ the random number) advancement in time. For the boundary spins the criterion in Algorithm 6 is applied. To ensure freedom of a deadlock a barrier is used for the boundary spins with a *wait until* the local time t becomes less than or equal to the same quantity for the neighbours.

For the kinetic Monte Carlo algorithm for the Ising model Lubachevsky [7.31] introduced an additional class N_b on top of the 10 classes for the boundary spins. Assume as above that the linear system size is L and that there are $4l$ boundary spins per processor. Then $N_\mathrm{b} = 4(l-1)$. The basic idea is to use the original Monte Carlo, for example Metropolis Monte Carlo, for the boundary spins and for the interior spins the kinetic Monte Carlo. Thus the algorithms proceeds as outlined in Algorithm 7. For this we augment the 10 classes with the additional class N_b.

Algorithm 7 Lubachevsky Parallel Kinetic Monte Carlo Algorithm

```
Initial time t = 0
for n-of-samples do
    Set up list of transition rates r_i = n_i p_i plus N_b
    Compute R_k = sum_{i=1}^{k} r_i
    Generate ρ from a uniform distribution on (0, 1]
    Choose i such that R_{i-1} < ρ R_i ≤ R_{i+1}
    Choose a spin with equal probability within the class i
    if spin is within the interior then
        Flip the spin
    else
        Wait until the local simulated time ≤ neighbour processor
        Apply Metropolis Monte Carlo to the spin
    end if
    Update time
end for
```

A slightly different approach has been taken by Martinez [7.44] by a synchronous time decomposition of the master equation (synchronous parallel kMC method (spkMC)). The basic idea is to create so called null events advancing the internal clock of each processor. This is done without altering the stochastic trajectory of the system. Further developments have been done specifically for the reaction-diffusion problems (see [7.45] and references therein).

Due to the success of other parallelization algorithms on GPUs, an algorithm was proposed by Jimenez and Ortiz [7.46], Klingbeil [7.47] and Agostino et al. [7.48]. Also discrete-event approaches have been developed [7.49] specifically for the Gillespie ansatz.

7.4 Lifting

As the name of this section suggests we augment the state space Ω with one or more additional variables. Let us first examine this idea for the Ising model in the case of conserved energy. Assume that we add the extra variable or degree of freedom to the Hamiltonian [7.50] (7.13)

$$\mathcal{H}' = e - \sum_{<i,j>} S_i S_j \quad S_i = \pm 1 \tag{7.19}$$

with $\Omega' = \mathbb{N} \times \Omega$. The extra variable e allows to lift the system out of the otherwise constraint hyperspace of constant energy. Set e to an appropriate value according to the initial energy. We can construct a Markov chain by choosing a spin at site ν at random. We change the spin direction at site ν to obtain $\Delta\mathcal{H}$ for the energy change in the Ising Hamiltonian. If we loose energy, then we transfer the energy to e and accept the change. If we would gain energy, then we accept the change under the condition that e has enough energy.

Let us now look at the more general case. Chen et al. and others [7.51–7.55] constructed a non-reversible Markov chain Monte Carlo Method (Lifted Metropolis-Hastings) as for example also in the (Hamiltonian) Hybrid Monte [7.56] (see also for the Bouncy Particle Sampler method [7.57]). The effect of this lifting is a reduced mixing time of the Markov chain (at best reduced by the square root of the original time).

So far we almost always used the detailed balance condition for the transition probability W and the invariant distribution p which we want to obtain from a Markov chain

$$p(x)W(x,x') = p(x')W(x',x) \quad \text{for all } x, x' \in \Omega\,. \tag{7.20}$$

This is not a necessary but sufficient a condition for the transition probability. One of possible solutions to (7.20) is the Metropolis Hastings transition probability

$$W(x,x') = q(x|x') \min\left\{1, \frac{p(x')q(x|x')}{q(x'|x)p(x)}\right\} \tag{7.21}$$

with the propositional probability q. The Hybrid Monte Carlo Method [7.56] has made use of this propositional probability.

Consider the global balance condition for the transition probability W [7.58]

$$\int p(x)W(x,x')dx = \int p(x')W(x',x)dx' \tag{7.22}$$

which we need to really to fulfill and the constraint

$$W(x,x')W(x',x) = 0 \text{ for all } x, x' \in \Omega \,. \tag{7.23}$$

W's that fulfill criterion 7.22 and criterion 7.23 are said to check a maximal global balance condition [7.59, 7.60].

Following the idea of adding additional degrees of freedom, we augment the system by an auxiliary variable e. Thus for the distribution p this results in

$$p(x,e) = p(x)p(e) \tag{7.24}$$

and for the above example (7.19) this would be

$$p(e) \propto \exp\{-\beta e\} \,. \tag{7.25}$$

and fix the propositional probability as

$$q(x',e|x,e) = \begin{cases} 1, & \text{if } x' = x + e\Delta s \\ 0, & \text{otherwise} \end{cases} \tag{7.26}$$

where the statement $x' = x + e\Delta s$ is meant to express that x' and x should not differ too much. Thus we updated the state in the direction given by e. This is continued until rejection occurs. Then we choose a new e' and continue with (x', e') which lifts the rejection into the lifting space rendering the entire method rejection-free. The probability for the choice of e' is based on the condition (7.22).

7.5 Event-Chain Monte Carlo

We will extend the rejection-free Monte Carlo simulation methods by considering irreversible Markov chains drawing on idea by Peters [7.61] and the concept of lifting [7.52]. These methods have been successfully developed for the problem of melting in two dimensions [7.62–7.65]. Extensions have been derived for discrete-variable models [7.66], classical continuous spin models [7.67, 7.68] and further generalized to rejection-free global-balance algorithms [7.69] and the forward event-chain Monte Carlo algorithm [7.58].

Here we follow [7.58] in the exposition of the algorithm. The goal is to use the ideas of lifting developed in the previous section to develop a rejection-free Monte Carlo algorithm. We use the extra variable e to suggest a new state. Rather than using an except/reject on this, we choose a time for the new event to happen and sample all the state in a chain along the way, until we have reached the transition time. We

then choose a new e value and continue. To sample the time Δs we go about as in (7.9) and (7.15). For ease presentation we follow the mechanistic language and assume an energy function $E(x)$ and consider e a velocity (see also Hybrid Monte Carlo [7.56]). Thus in (7.26) we are looking for displacements in space controlled by the time Δs and the velocity e.

In (7.26) we have made a choice for the propositional probability. With the notation $[a]^+ = \max\{0, a\}$ and Metropolis choice of transition probability (7.21) we have

$$W(x, x') = \min\{1, \exp\{-\Delta E(x)e\}\} = \exp\{-[\Delta E(x)e]^+\} \,. \tag{7.27}$$

To determine the transition time we add up all the moves until we have reached the event time

$$\Delta E^*(\Delta s) = \int_0^{\Delta s} [\Delta E(x + se)e]^+ \mathrm{d}s \tag{7.28}$$

and find the time Δs by solving the equation

$$\Delta E^*(\Delta s) = \log(\rho) \tag{7.29}$$

where $\rho \in (0, 1]$ is a uniform random number. It rests to choose the transition probability for e. Here Michel and Senecal [7.58] suggest

$$p(e' \to e) = \delta(e' + e) \tag{7.30}$$

In Algorithm 8 the full algorithm is exposed (for parallelization for example for dense hard sphere and polymer systems see [7.70]).

Algorithm 8 Event Chain Monte Carlo Algorithm [7.58]

Initial state $x' = x_0$
for n-of-samples **do**
  Set current event chain length $l_c = l$
  Set random direction e
  while True **do**
    Set initial sample $x = x'$
    Compute $\Delta E^* = -\log(\rho)$, ρ from a uniform distribution on (0, 1]
    Compute Δs
    if $l_c < \Delta s$ **then**
      Compute $x' = x + l_c e$
      Set sample $x^k = x'$
      Break
    else
      Compute $x' = x + \Delta s e$
      Update chain length $l_c = l_c - \Delta s$
      Update direction $-e$
    end if
  end while
end for

Chapter 8
Finite Size Scaling Tools for the Study of Interfacial Phenomena and Wetting

8.1 Introduction

In this chapter, we use the word "interface" in the sense of a boundary between coexisting bulk phases (in thermal equilibrium). An example is the interface between liquid (e.g. water) and gas phases (water vapor) but also interfaces between fluid and solid phases (e.g. water and ice) can be considered, as well as interfaces between coexisting solid phases. The generic example are "domain walls" in magnets, separating domains with opposite orientation of the magnetization, a case that can already be studied in the framework of the simple Ising model (Chaps. 2 and 3) where one has spins on the sites of a rigid perfect lattice pointing up or down.

A crucial feature of such interfaces is that statistical fluctuations are possible both in the bulk phases (more or less far away from the separating interface) and also with respect to the interface itself (e.g. its local position relative to a reference dividing plane, in $d = 3$ dimensions). In fact, in many cases interfaces can be interpreted as mesoscopic objects with degrees of freedom on many length scales (below we shall consider the effect of long wavelength fluctuations, the so-called "capillary waves"). For this reason, the finiteness of the linear dimension of interfaces (e.g. measured via the area of the above reference dividing plane) is a crucial aspect, causing important finite size effects which must be considered when one wishes to study interfaces (at nonzero temperature) by computer simulation. A crucial observation also is that directions parallel and perpendicular to the reference plane are not equivalent; this anisotropy of the interfacial fluctuations needs to be taken into account when one tries to handle the associated finite size effects in terms of appropriate extensions of finite size scaling theories (Sects. 8.3 and 8.4).

The excess free energy due to an interface, the so-called "interfacial tension", in the literature often also is called "surface tension". However, in the present chapter we wish to use the word "surface" in a different meaning: e.g., a magnetic crystal (described, for instance, by an Ising model) may have a surface plane against vacuum, such that there are no spins on the other side of the surface plane (this case was already mentioned in the context of the "free boundary condition", see Sect. 2.2.2). For a

K. Binder and D. W. Heermann, *Monte Carlo Simulation in Statistical Physics*,
Graduate Texts in Physics, https://doi.org/10.1007/978-3-030-10758-1_8

fluid system, we may consider the solid wall of a container as such a surface which simply acts as a boundary condition in terms of a suitable potential. This potential acts on the fluid particles, preventing them also from crossing the surface (and thus particles cannot leave the container). Depending on the physical situation envisaged, the surface potential may be independent of the x,y-coordinates in the surface plane (representing a flat structureless surface) or not. The latter case may represent just the so-called "corrugation" of a wall due to the regular periodic arrangement of atoms in a crystal structure of a solid, or a nanoscopic/mesoscopic inhomogeneity, as sometimes is of interest in the case of nanostructured materials. Fluids near the structured walls indeed have become a very active research topic in this century, and Monte Carlo simulation is a valuable tool to study the properties of fluids under such conditions. Again, such boundaries cause an excess free energy, which often is called "surface tension" also but will be called "wall tension" here to avoid confusion with the interfacial tension.

Estimation of interfacial and wall tensions by Monte Carlo requires nontrivial extensions of the methodology described in earlier chapters of this book, and shall hence be considered in the present chapter.

A very important facet of these problems occurs when a fluid confined by walls may exhibit coexistence between two phases separated by an interface: the potential acting on the fluid particles due to a wall may lead to an effective "interfacial potential" exerted by the wall on an interface. This potential may be repulsive or attractive, and in general will exhibit a nontrivial dependence on macroscopic control parameters (such as temperature or pressure of the system, etc). In such systems, one may observe the so-called "wetting transitions", which may also be considered as "unbinding transitions" of a bound interface from the considered wall [8.1–8.4]. Just as the liquid-vapor transition in the bulk is a singularity of the bulk free energy of the fluid, is a wetting transition (which may be of first or second order) a singularity of the wall tension of the considered fluid. As has been emphasized throughout this book, finite systems do not show the singularities associated with phase transitions, rather these singularities are rounded (and often also shifted) due to finite size. Computer simulations deal with finite systems exclusively, and hence finite size effects need consideration when one wishes to characterize wetting phenomena quantitatively (Sect. 8.3).

As a final disclaimer, we stress that this chapter is intended as a tutorial guide for the newcomer, and not at all as a review article that would cover all the work that exists in the field (and is quite diverse and abundant). When examples are taken from the groups of the authors of this book, this is done for the sake of convenience only, and does not mean that related work of other groups lacks value.

8.2 A Reminder on Finite Size Scaling of the Order Parameter at Bulk 1st and 2nd Order Phase Transitions

8.2.1 First Order Transitions

At a thermally driven first order transition, such as melting of a crystalline solid, or the order-disorder transition of the q-state Potts model [8.5, 8.6] on a lattice that was already discussed in Chap. 2, the low temperature phase is characterized by an order parameter which (in the thermodynamic limit) discontinuously vanishes at the transition temperature T_c. The free energy $F(T)$ near T_c varies linearly as a function of temperature,

$$F_-(T < T_c) = F_c - S_-(T - T_c), \tag{8.1a}$$

$$F_+(T > T_c) = F_c - S_+(T - T_c), \tag{8.1b}$$

where F_c is the free energy right at T_c (where the free energy branches of the ordered ($T < T_c$) and disordered ($T > T_c$) phases cross), and S_-, S_+ are the entropies at the transition (both entropy S and internal energy E exhibit discontinuous jumps at T_c, in the thermodynamic limit, particle number $N \to \infty$). Since $F = E - TS$, we can express the difference $\Delta F = F_- - F_+$ near T_c also in terms of the jump $\Delta E = E_+ - E_-$ of the internal energy,

$$\Delta F = \Delta E(T - T_c)/T_c \quad . \tag{8.2}$$

It is convenient (in a lattice model, assuming a cubic lattice of linear dimension L with periodic boundary conditions in all d directions of the d-dimensional system) to define the above quantities F, S, E per lattice site (for off-lattice systems, we could use a normalization per unit volume). Then the statistical weights of both phases near the transition are proportional to the quantities a_+, a_- introduced already in Chap. 2,

$$a_+ = q_+ \exp(\Delta F L^d / 2k_B T), \; a_- = q_- \exp(-\Delta F L^d / 2k_B T), \tag{8.3}$$

where q_-, q_+ are the degeneracies of the phases. For the q-state Potts model, the disordered phase is non-degenerate, $q_+ \equiv 1$, while $q_- = q$, since there are q different domains the system can be in; but there exist also cases where one encounters a first-order phase transition between two phases exhibiting a different kind of order. An example is the anisotropic Heisenberg antiferromagnet in a uniform magnetic field H, where for H less than H_t one has an antiferromagnet of Ising type ($q_+ = 2$) and for $H > H_t$ a spin-flop phase (with an XY-like type of order) [8.7–8.9]. The extension of the finite size scaling description [8.10, 8.11] that is outlined below to this field-driven transition is given elsewhere [8.12]. Using $\Delta T = T - T_c$ we can rewrite (8.2 and 8.3) as

$$a_+ = q_+ \exp(\Delta T \Delta E L^d / 2k_B T T_c), \; a_- = q_-(-\Delta T \Delta E L^d / 2k_B T T_c) \tag{8.4}$$

which was already used in Chap. 2 to discuss the finite size rounding of the specific heat $C(T, L) = \partial \langle E \rangle_L / \partial T$ at the transition.

Here we shall rather focus on the description of the distribution of the order parameter $\vec{\psi}$ (recall that for the Potts model the order parameter has q discrete orientations in a $(q-1)$ dimensional vector space [8.6]). In the finite system, for $T > T_c$ and large enough L the distribution simply is a multivariate Gaussian distribution,

$$P_L^+(\vec{\psi}) = \mathcal{N} \exp\Big(- \frac{\vec{\psi}^2 L^d}{2 k_B T \tilde{\chi}^+} \Big), \tag{8.5}$$

where $\tilde{\chi}^+$ is a kind of "susceptibility" measuring the strength of the order parameter fluctuations in the disordered phase, and $\mathcal{N}$ is a normalization factor. Likewise, for $T < T_c$ the distribution is a sum of q Gaussians centered at the discrete values of the order parameter $\vec{\psi}_k$ $(k = 1, \ldots, q)$ representing the q possible ordered states. Clearly, this is an obvious generalization of the order parameter distribution of the Ising model (with a scalar, one-component order parameter) discussed in Chap. 2. Here we rather point to the fact that this description already yields an interesting consequence as the thermodynamic limit is approached: the relative weight of the disordered phase is $a_+/(a_+ + a_-) = 1$ for $T > T_c$ and 0 for $T < T_c$ but takes a nontrivial value $q_+/(q_+ + q_-)$ for $T = T_c$. Hence for $L \to \infty$ and $T = T_c$ we can write the order parameter distribution as a sum of δ-functions as follows, for $T = T_c$ (recalling $q_+ = 1, q_- = q$)

$$P_\infty(\vec{\psi}) \propto \delta(\vec{\psi}) + \sum_{k=1}^{q} \delta(\vec{\psi} - \vec{\psi}_k), \tag{8.6}$$

with $\langle |\vec{\psi}_k| \rangle_\infty = \langle \vec{\psi}_k^2 \rangle_\infty^{1/2} = \psi_\infty$ being the value of the order parameter in the limit $T \to T_c^-$ in the thermodynamic limit. From (8.6) one readily finds that at $T = T_c$

$$\langle \psi^2 \rangle_\infty = \psi_\infty^2 q/(1+q), \ \langle \psi^4 \rangle_\infty = \psi_\infty^4 q/(1+q) \tag{8.7}$$

and hence the fourth-order cumulant [8.13] $U_L^\psi = 1 - \langle \psi^4 \rangle_L / [3 \langle \psi^2 \rangle_L^2]$ takes a nontrivial value at T_c also at a first-order transition,

$$U_\infty^\psi = 1 - \langle \psi^4 \rangle_\infty / [3 \langle \psi^2 \rangle_\infty^2] = 2/3 - 1/(3q), \quad T = T_c \,. \tag{8.8}$$

Thus, U_∞^ψ stays 2/3 for all $T < T_c$, jumps at T_c to U_∞^ψ as quoted in (8.8), and jumps to zero for $T > T_c$. For large but finite L, when the delta functions in (8.6) are replaced by the appropriate Gaussians (such as (8.5)), the quantities U_L^ψ become smooth functions of T, but still intersect at an (almost) common intersection point, given by (8.8). Finite-size corrections to this common intersection point are small of order L^{-d}, while a shift of this intersection point away from T_c is small of order L^{-2d} [8.11]. As is obvious from (8.4), the width of the region over which the jump of U_L^ψ from 2/3 to U_∞^ψ and further to zero is rounded also is of order $\Delta T \propto T_c/(\Delta E L^d)$.

However, a nontrivial feature is the prediction that U_L^{ψ} exhibits a minimum whose depth diverges to minus infinity proportional to L^d at a position $T_{\min} - T_c \propto L^{-d}$ above T_c [8.11]. This prediction of the phenomenological scaling theory has been verified by Monte Carlo simulations for the 3-state Potts model in $d = 3$ [8.11]. Ideas related to the above description were also developed in [8.14, 8.15].

Thus the observation of a common order parameter cumulant intersection point is not always a reliable evidence for a second-order transition; if the cumulant exhibits a deep minimum, rather a first-order transition can be expected. However, as a general caveat we mention that the correct sampling of the relative weights of the various phases (that can coexist right at T_c) in the region around T_c often is a very nontrivial problem, requiring a huge numerical effort. Methods which have been advocated for this purpose are the replica-exchange framework for Wang-Landau sampling [8.16] and "simulated tempering" [8.17] with approximated weights [8.18], for instance. For a general introduction into these (and other) methods to efficiently sample free energy "landscapes" accurately, we refer the reader to Chap. 6 in this book. We emphasize, however, that distinguishing correctly the order of phase transitions (and locating accurately at which values of the control parameter they occur) is not just an academic problem of 20th century physics, but still relevant in many contexts, e.g. phase transitions discussed in quantum chromodynamics (QCD) [8.19, 8.20]. A recent overview of QCD as a theory of strong interactions between elementary particles in the framework of lattice gauge theory and the role of Monte Carlo simulations of QCD on high performance computing can be found in [8.21] but this subject is out of our scope here; we only mention that the phase transition of interest there is the "condensation" of hadrons from the quark-gluon plasma, a phenomenon that is thought to have occurred in the evolution of the early universe.

As a final remark on first-order transitions, we emphasize that a useful consequence of (8.5)–(8.7) is that not only the fourth-order cumulant U_L^{ψ} of the order parameter intersects at the common intersection point U_{∞}^{ψ} for large L, but actually such common intersection points do occur for all moments of the order parameter, and the values of these intersections for $\langle \psi^2 \rangle_L$ and $\langle \psi^4 \rangle_L$ are simply the expressions quoted in (8.7). This property is different for a bulk second-order transition, where moments do not have common intersection points, as we shall see below. There is also a simple scaling property [8.12]

$$\langle \psi^2 \rangle_L / \psi_{\infty}^2 = [b + q \exp(\mathcal{Z})]/[1 + q \exp(\mathcal{Z}] \tag{8.9}$$

where b is a correction to scaling (of order L^{-d}) and $\mathcal{Z} = -\Delta F L^d / k_B T_c = -\Delta E \Delta T L^d / k_B T_c^2$. As demonstrated already in Chap. 2 for the specific heat, at first-order transitions scaling functions can be computed explicitly (in the framework of the Gaussian approximation for the distribution functions of pure phases).

8.2.2 Second Order Transitions

For simplicity, we consider now only the case where the order parameter is a scalar, s (such as in the Ising ferromagnet, discussed in Chap. 2). Equation (8.5) then gets replaced by

$$P_L^+(s) = \mathcal{N} \exp\Big(- \frac{s^2 L^d}{2k_B T \chi^+}\Big), \quad L \gg \xi \quad , \tag{8.10}$$

noting that L must (by far) exceed the correlation length ξ. Here $\chi^+(\chi^-)$ is the ferromagnetic susceptibility, which diverges when one approaches T_c, $\chi^+ = \hat{\Gamma}^+\Big(\frac{T}{T_c} - 1\Big)^{-\gamma}$, $\chi^- = \hat{\Gamma}^-\Big(1 - T/T_c\Big)^{-\gamma}$, while $\xi = \hat{\xi}_\pm |1 - T/T_c|^{-\nu}$; γ and ν are the standard critical exponents of susceptibility and correlation length, respectively [8.22, 8.23]. $\hat{\Gamma}^+$, $\hat{\Gamma}^-$ and $\hat{\xi}$, $\hat{\xi}^-$ are the associated critical amplitudes [8.23].

For $T < T_c$ and $L \gg \xi$, the system is in the ordered phase, with $|s|$ near the bulk order parameter s_∞, with exponent β [$s_\infty = \hat{B}(1 - T/T_c)^\beta$]. We recall from Chap. 2 the double-Gaussian approximation, for $T < T_c$,

$$P_L^-(s) \propto \exp\Big[- \frac{(s - s_\infty)^2 L^d}{2k_B T \chi^-}\Big] + \exp\Big[- \frac{(s + s_\infty)^2 L^d}{2k_B T \chi^-}\Big] \quad , \tag{8.11}$$

which should hold for s near $\pm s_\infty$ only, of course. Rescaling s by s_∞, $\tilde{s} = s/s_\infty$, (8.11) can be written as, scaling L by ξ as well ($\tilde{L} = L/\xi$)

$$P_L^-\Big(\tilde{s}, \tilde{L}\Big) \propto \exp\Big[- \frac{(\tilde{s} - 1)^2 \hat{\xi}^{-d}}{2k_B T \hat{\Gamma}^-/\hat{B}^2} \tilde{L}^d \Big(1 - T/T_c\Big)^{2\beta+\gamma-d\nu}\Big]$$
$$+ \exp\Big[- \frac{(\tilde{s} + 1)^2 \hat{\xi}^{-d}}{2k_B T \hat{\Gamma}^-/\hat{B}^2} \tilde{L}^d \Big(1 - T/T_c\Big)^{2\beta+\gamma-d\nu}\Big] \quad . \tag{8.12}$$

Now if the "hyperscaling relation"

$$d\nu = \gamma + 2\beta \tag{8.13}$$

among the critical exponents holds, the temperature dependence has completely disappeared from $P_L^-(\tilde{s}, \tilde{L})$. From the two scaling variables, $\tilde{s}$, $\tilde{L}$ we can construct another scaling variable $sL^{\beta/\nu}$ and hence we see that the order parameter distribution $P_L^-(s)$ [but this holds also for $P_L^+(s)$] can be considered as a function of the two variables $sL^{\beta/\nu}$, L/ξ, i.e. [8.13]

$$P_L(s) = L^{\beta/\nu} \tilde{p}(sL^{\beta/\nu}, L/\xi) \quad , L \to \infty \quad . \tag{8.14}$$

While (8.10) and (8.12) hold for $L/\xi \gg 1$ only, (8.14) is supposed to hold for all ratios L/ξ. From (8.14) it is straightforwardly seen that

$$\langle s^{2k} \rangle = L^{-2k\beta/\nu} \tilde{s}_{2k}(L/\xi) \quad , k = 1, 2, \ldots \quad . \tag{8.15}$$

Thus, at T_c (where $\xi = \infty$) due to the power law prefactor $L^{-2\beta/\nu}$ there is no L-independent common intersection point $\langle s^2 \rangle$ here, unlike the first-order case. Finally the cumulant becomes

$$U_L = 1 - \langle s^4 \rangle / [3 \langle s^2 \rangle^2] = \tilde{U}(L/\xi) \quad , \tag{8.16}$$

as discussed in more detail in Chap. 2 already. Here we like to draw attention to the case of the random field Ising model (RFIM) [8.24–8.26]. In this model one has a random field which is zero on average and spatially uncorrelated (either $\pm h_r$ or drawn from a Gaussian distribution, for instance), to model some quenched random disorder, in the system. For this model it is known that (8.13) does not hold, and rather is replaced by

$$\gamma + 2\beta = \nu(d - \theta) \quad , \quad \theta = \gamma/\nu; \tag{8.17}$$

the exponent θ characterizes the violation of hyperscaling. From (8.12) and (8.17) we immediately recognize that in the argument of the exponential functions a factor $(1 - T/T_c)^{-\nu\theta}$ must remain: i.e., in the scaling limit ($\tilde{s}$ near unity, $\tilde{L}$ finite, $T \to T_c$), the order parameter distribution tends towards a sum of two delta functions (and also $P_L^+(\tilde{s}, \tilde{L})$ tends to a delta function centered at $\tilde{s} = 0$). As a consequence, in this model there is no common cumulant intersection point, as found also by explicit Monte Carlo simulation results for the RFIM and various related models [8.27–8.30]. For this reason, the estimation of both T_c and the critical exponents of the RFIM (which differ from those of the pure Ising model, of course) is a subtle task [8.30]. We also recall that $d = 2$ is the lower critical dimension for the RFIM [8.24, 8.26, 8.31]: long-range ferromagnetic order is unstable in $d = 2$, since the system can spontaneously break up into domains of finite size, since there exists a length scale $L_{\text{domain}}(h_r)$ for which the interface tension vanishes [8.31].

We now summarize the results on the finite size effects of the order parameter distributions at phase transitions as follows: for a system of finite linear dimension L (in a (hyper) cubic geometry, with periodic boundary conditions) all moments of the order parameter distribution ($\langle \psi^2 \rangle_L$, $\langle (\psi^2)^2 \rangle_L \ldots$) are smooth nonsingular functions of the control parameter (e.g.temperature T). At first-order transitions, the transition temperature T_c shows up as a common intersection point of both these moments and the fourth-order cumulant (cf. (8.7) and (8.8)). Scaling functions (e.g. (8.9)) can be simply computed, if pure phases are described in terms of Gaussians (e.g. (8.5)). At second-order transition (with a nonzero order parameter exponent β) moments at T_c do not intersect at a common intersection point, but rather vary as the appropriate power of $L^{-\beta/\nu}$ (8.15). Scaling functions such as $\tilde{s}_{2k}(L/\xi)$, $\tilde{U}(L/\xi)$ are nontrivial. Cumulants (8.16) do have (for $L \to \infty$, where corrections to finite size scaling are negligible, such that (8.14) holds) a common intersection point, which hence is useful to locate T_c. However, the value of $\tilde{U}(0)$ at which this intersection occurs

is a nontrivial characteristic of the "universality class" of the considered critical phenomenon (just like critical exponents [8.22], critical amplitude ratios [8.23], etc.) All these statements presuppose the validity of the hyperscaling relation, (8.13), which normally can be taken for granted. Only in special cases (e.g. the RFIM) where hyperscaling is violated no unique cumulant intersection is possible, since in this case (8.17) the order parameter distribution (e.g. (8.13)) cannot be reduced to the simple scaling form, (8.14), it rather tends to several δ-function peaks. Note that another violation of hyperscaling occurs for systems with mean-field critical exponents (see Chap. 2); in the latter case there is a simple scaling similar to (8.14), but the correlation length needs to be replaced by a "thermodynamic length" $\ell_T = (k_B T \chi \psi^{-2})^{1/d}$ [8.32].

8.2.3 *Anisotropic Finite Size Scaling and Its Application to Wetting Phenomena: A "Crash Course"*

For the benefit of the non-specialist reader we summarize in this subsection the essential concepts about wetting phenomena and their physical content, as a prelude to the explanations of the problems encountered by simulation studies of wetting. Let us consider the vapor to liquid transition of a fluid, focusing first on the situation where a saturated gas is exposed to a solid wall, that exerts attractive forces to the particles in the fluid (Fig. 8.1a). If the wall is not wetted by the fluid, a (macroscopically large) liquid droplet attached to the wall would exhibit a nonzero contact angle θ. As is well known [8.1–8.4], the magnitude of this contact angle is controlled by the competition of the liquid-gas interfacial tension ($\sigma_{\ell g}$) with the difference in wall tensions of the two coexisting gas (σ_{wg}) and liquid ($\sigma_{w\ell}$) phases, namely by Young's equation.

$$\cos\theta = (\sigma_{wg} - \sigma_{w\ell})/\sigma_{\ell g} \quad , \quad \text{if } \sigma_{\ell g} \geq |\sigma_{wg} - \sigma_{w\ell}| \quad . \tag{8.18}$$

If the inequality is not fulfilled, we have either complete wetting of the wall (if $\sigma_{wg} - \sigma_{w\ell} > \sigma_{\ell g}$, the actual surface excess energy of the vapor phase then is given by $\sigma_{wg} \equiv \sigma_{w\ell} + \sigma_{\ell g}$, because then the wall is coated by a macroscopically thick liquid layer), or we have complete drying of the liquid phase (if $\sigma_{w\ell} - \sigma_{wg} > \sigma_{\ell g}$, the actual surface excess free energy of the liquid phase is given by $\sigma_{w\ell} = \sigma_{wg} + \sigma_{\ell g}$, because then the liquid does not extend up to the wall, which is then coated by a macroscopically thick vapor layer). In the framework of the lattice gas (Ising) model of fluids, that is emphasized in this chapter, the particle hole/symmetry of the model also implies a symmetry between wetting and drying phenomena, just the roles of liquid and vapor are interchanged, and hence we shall not discuss "drying" ($\theta = 180°$; droplets of the liquid would then energetically prefer not to touch the wall) further. However, Monte Carlo simulations of wetting/drying phenomena can also be performed [8.33, 8.34] for models that lack this symmetry between coexisting phases, e.g. the Asakura-Oosawa model [8.35–8.37] of colloid-polymer mixtures.

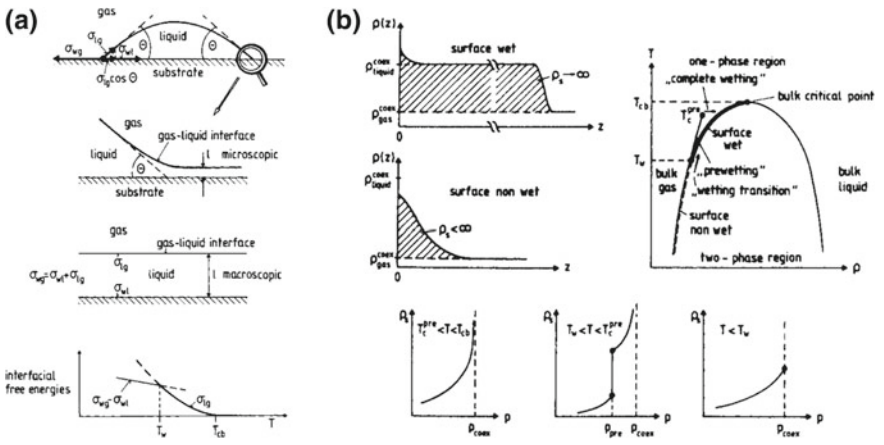


Fig. 8.1 **a** Schematic cross section through a (macroscopic) droplet of liquid, on a substrate surface for the case of partial wetting, showing the contact angle θ and the interpretation of Young's equation [8.1–8.4] in terms of mechanical equilibrium of the surface tensions at the contact line (upper part). A magnified view of the three phase contact line, where the liquid-gas interface meets the substrate surface, is shown just below the topmost part. Note that this picture is still coarse-grained, the cross section of the liquid gas interface is just shown as a line (neglecting fluctuations, intrinsic interfacial thickness, etc.). On a scale of (at least) a few molecular diameters, this contact line is rounded away, since the liquid-gas interface bends over into a flat liquid film of microscopically small thickness ℓ (of the order of a few molecular diameters). In contrast, for the case of complete wetting of the surface (middle part), the substrate is covered by a uniform liquid film of a macroscopic thickness ℓ, and droplets cannot exist and spread out ($\theta = 0$). Due to the temperature dependence of the various interfacial free energies, a wetting transition at a temperature $T = T_w$ may occur (bottom part), the substrate being "dry" (or partially wet) for $T < T_w$ and wet for $T > T_w$. **b** Schematic phase diagram of a semi-infinite fluid in contact with a wall (upper right part) showing in the temperature (T) density (ρ) plane the coexistence curve in the bulk. The enclosed two-phase coexistence region ends at the bulk critical point (T_{cb}). The wetting transition is a singularity of the fluid associated with the wall and occurs at the temperature T_w at the vapor density according to the gas branch of the coexistence curve. The corresponding profile of the (coarse-grained) density $\rho(z)$ of the liquid (z is the distance from the wall) from a partially wet (or non-wet) state of the surface (left middle part of the figure) and a completely wet state (left upper part of the figure). The area of the shaded region in these plots denotes the surface excess density ρ_s. Along the coexistence curve ρ_s is finite for $T < T_w$ and jumps to infinity at T_w, for a first-order wetting transition. The lower part of the figure shows the variation of ρ_s with ρ in the undersaturated vapor region. For $T_w < T < T_c^{pre}$ (the prewetting critical point) there occurs at ρ_{pre} a first-order prewetting transition (jump of ρ_s). From Binder et al. [8.3]

As it has already been emphasized above, this description of wetting in terms of the contact angle of droplets really is a macroscopic description. When it is attempted to directly simulate wall-attached droplets (e.g. [8.38–8.48], one encounters numerous difficulties: (i) for a droplet of finite volume, the vapor pressure of the fluid surrounding the droplet in equilibrium is enhanced in comparison with the bulk coexistence pressure (Gibbs-Thomson effect). Since the surface tension of the droplet depends on its radius of curvature, analysis of the equilibrium between the droplet and the

vapor is very subtle [8.42–8.48]. A simplification only arises for "nonvolatile liquids" (e.g., polymer chains), for which the vapor density can be strictly assumed to be zero (e.g. [8.40]). (ii) Small droplets exhibit very long-lived fluctuations of their shape and surface area, and hence it is extremely difficult to obtain statistically significant data (see again [8.40] for explicit examples of such fluctuations). (iii) The "contact line" where the gas-liquid interface of the droplet meets the wall causes another free energy excess contribution, the so-called "line tension" [8.49–8.51], which modifies the contact angle of a droplet of radius R by a $1/R$ correction [8.52–8.54].

Of course, all these difficulties are avoided when one takes the Young equation, (8.18), for granted, and just tries to find the relevant interfacial and surface excess free energies ($\sigma_{\ell g}$, σ_{wg}, $\sigma_{w\ell}$) from suitable simulations separately. Recording the temperature dependence of both the liquid-gas interfacial tension $\sigma_{\ell g}(T)$ and of the difference of the wall-fluid surface tensions of the gas and the liquid, $\sigma_{wg}(T) - \sigma_{w\ell}(T)$, cf. Fig. 8.1a, one can locate the wetting transition temperature. Section 8.4 will describe briefly some of the methods available for the estimation of such interfacial tensions from Monte Carlo methods.

We emphasize that the intersection of two branches of interfacial free energies (Fig. 8.1a bottom part) with different slopes at T_w corresponds to a first order wetting transition. This is the analog of (8.1a), (8.1b) and (8.2), for surface excess free energies rather than bulk free energies.

However, just as in the bulk where both first-order and second-order transitions are known to occur, also wetting transitions of both first- and second-order can occur. For a second-order transition ("critical wetting" [8.1–8.4]) the two interfacial free energy branches in the bottom part of Fig. 8.1a meet with a common tangent at T_w. Of course, then the estimation of the interfacial free energies by simulation methods is not a useful route to locate T_w, in particular since also at a second-order wetting transition critical fluctuations occur: in the thermodynamic limit, critical wetting is characterized by both divergent correlation lengths and by divergent relaxation times ("critical slowing down" [8.55]).

The diverging correlation lengths at critical wetting call for an extension of finite size scaling methods to this case. However, this problem is not straightforward, since two distinct correlation lengths occur, characterizing the decay of fluctuations in different directions: parallel ($\xi_{||}$) and perpendicular ($\xi_{\perp}$) to the wall. Note that the (coarse-grained) density $\rho(z)$ in the z-direction normal to the attractive wall (Fig. 8.1b) becomes gradually broader as $T \to T_w$ from below, and exhibits at T slightly below T_w already a two-step decay, similar to the completely wet case: $\rho(z)$ first decreases only up to the value $\rho_{\text{liquid}}^{\text{coex}}$, the liquid density at coexistence, and only at a much larger distance ℓ from the wall the second step follows, where $\rho(z)$ decays from $\rho_{\text{liquid}}^{\text{coex}}$ to $\rho_{\text{gas}}^{\text{coex}}$ (Fig. 8.1b). This second step is nothing but the ordinary liquid-vapor interfacial profile of the density, when ℓ is large enough. Critical wetting hence can be interpreted as gradual unbinding of an interface from a wall that exhibits an effective attractive interaction with the interface. The surface excess density ρ_s then does not reach a finite constant as it would for a 1st order wetting transition (Fig. 8.1b) but rather shows a critical divergence as $T \to T_w$. Note that for critical wetting the prewetting phenomenon included in Fig. 8.1b does not occur.

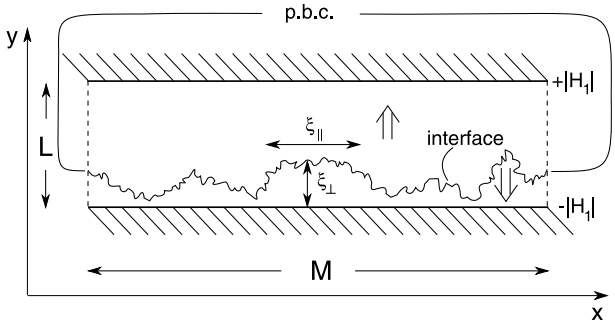


Fig. 8.2 Schematic description of the system geometry used for simulations of wetting transitions in Ising models and the occurring fluctuations of an interface bound to the lower wall. For simplicity, a two-dimensional system is shown, using a $M \times L$ lattice, with periodic boundary conditions in x-direction, while at $y = 1$ and $y = L$ free boundaries are used, at which surface magnetic fields $\pm|H_1|$ act. The fluctuating interface is indicated, separating domains of positive and negative magnetization (the sign of the magnetization is indicated by the double arrow; it is assumed that a coarse-graining has been observed on a length scale intermediate between the correlation length ξ in the bulk and the length scale $\xi_\perp$ characterizing the extent of the interfacial fluctuations in the y-direction). Then fluctuations of the bulk magnetization in the interior of the domain are averaged over, and only interfacial fluctuations are left. The local excursions of the interface relative to its average position in the y-direction are correlated in the x-direction over a length scale $\xi_{||}$. From Albano and Binder [8.56]

Now it is important to realize that the profiles of $\rho(z)$ shown schematically in Fig. 8.1b refer to the thermal average, while the interface weakly bound to the wall is a strongly fluctuating object (Fig. 8.2) [8.56]. An instantaneous snapshot of the system in a Monte Carlo simulation then will resemble Fig. 8.2, if the bulk fluctuations inside the domains are averaged over. These interfacial fluctuations are strongly anisotropic, directions perpendicular to the interface and parallel to it are not equivalent, and there is the need to introduce two correlation lengths $\xi_{||}$, $\xi_\perp$, which diverge as $T \to T_w$ with different exponents

$$\xi_{||} \propto (T_w - T)^{-\nu_{||}}, \quad \xi_\perp \propto (T_w - T)^{-\nu_\perp} \quad . \tag{8.19}$$

Equation (8.19) can be shown rigorously to be true for the $d = 2$ semi-infinite Ising model with a surface field $-|H_1|$, where also the location of T_w can be exactly predicted [8.57–8.59] in terms of the inverse function $H_{1c}(T)$ of $T_w(|H_1|)$, for the square Ising lattice

$$\cosh(2H_{1c}/k_B T) = \cosh(2J/k_B T) - \sinh(2J/k_B T)\exp(-2J/k_B T) \tag{8.20}$$

and $\nu_{||} = 2$, $\nu_\perp = 1$. Mean-field theory of critical wetting (with short-range surface forces, as appropriate for an Ising model with a local surface magnetic field) implies $\nu_{||} = 1, \nu_\perp = 0$ (i.e., a logarithmic divergence), however [8.1–8.4]. While mean-field theory is believed to hold for $d > 3$, the case $d = 3$ is expected to show nontrivial

critical behavior [8.60] and has led to a longstanding debate (e.g. [8.61–8.67]) that still is unsettled.

Finally, we emphasize that strictly speaking for wetting phenomena we should consider a semi-infinite system (as anticipated in Fig. 8.1) while Fig. 8.2 rather considers a system which is finite in y-direction and uses there an antisymmetric boundary condition. Actually, early simulations of wetting phenomena [8.61–8.63] have not used this geometry, but rather used a "thick" film with symmetric walls, so that (in the incompletely wet state) weakly bound interfaces occur at both walls (one then has domains with negative magnetizations near both walls, (and a (thicker) domain with positive magnetization in the center of the film). However, the disadvantage of this simulation geometry is that without a positive field in the bulk the considered system is only metastable: the system lowers its free energy, when the two interfaces meet in the center and annihilate the domain with positive magnetization. If one would apply a positive bulk field to prevent this, one changes the character of the wetting phenomena significantly (it corresponds to working off coexistence in the phase diagram of Fig. 8.1b).

In the geometry of Fig. 8.2, however, there exists always a single interface only, and one expects that in the state of partial wetting it is either bound to the lower wall (at $y = 1$) or to the upper wall (at $y = L$), due to the antisymmetry of the chosen surface fields there occurs a bistable symmetry of the state. In contrast, when wetting (interface unbinding from the wall) occurs, the interface will occur at the maximum possible distance from both walls (on average), i.e. in the center of the film. The resulting interfacial transition in such a thin film geometry is called interface localization/delocalization transition [8.68–8.71]; in order to study it one must keep the perpendicular linear dimension L finite and consider the limit where the linear dimensions of the walls (M) tend to infinity. However, only recently it was realized that taking both linear dimensions L and M towards infinity is the appropriate limit to extract information on wetting phenomena [8.56]. This will be explained in more detail in the next subsection. Before this is done, we formulate the scaling theory of critical wetting for a semi-infinite system. As has been said above, we consider the surface excess free energy (abbreviating $\sigma_{wg} - \sigma_{w\ell} = f_s$) or its singular part $f_s^{(sing)}$ [8.2]

$$f_s^{(sing)}/k_B T = |t|^{2-\alpha_s} \tilde{F}_s(H|t|^{-\Delta_s}), t = 1 - T/T_w(H_1) \to 0 \quad , \tag{8.21}$$

where H is the bulk field and α_s, Δ_s are appropriate critical exponents, defined in analogy to critical behavior in the bulk $\{ f_s^{(sing)} = |\tau|^{2-\alpha_b} \tilde{F}_b(H|\tau|^{-\Delta_b}), \tau = 1 - T/T_c\}$ [8.22, 8.23] with $\tilde{F}_s$, $\tilde{F}_b$ appropriate scaling functions. The surface exponents in $d = 2$ then simply are [8.2, 8.57–8.59]

$$\alpha_s = 0, \Delta_s = 3 \quad . \tag{8.22}$$

As in the bulk, where magnetization and susceptibility are derived as 1st and 2nd derivative of the free energy with respect to the field H, we find surface excess

magnetization m_s and surface excess susceptibility χ_s here as derivatives of the surface excess free energy f_s, i.e.

$$m_s = -(\partial f_s^{(\mathrm{sing})}/\partial H)_T \propto t^{2-\alpha_s-\Delta_s} \equiv t^{\beta_s} \quad , \tag{8.23}$$

$$\chi_s = -(\partial^2 f_s^{(\mathrm{sing})}/\partial H^2)_T \propto t^{2-\alpha_s-2\Delta_s} \equiv t^{-\gamma_s} \quad . \tag{8.24}$$

From (8.22)–(8.24) we readily get the values of the associated critical exponents

$$\beta_s = -1 \quad , \quad \gamma_s = 4 \tag{8.25}$$

The surface excess magnetization m_s can also be interpreted as the integral of the magnetization profile $\int_0^\infty \mathrm{d}z[m(z) - m_b] = m_s$, where the bulk magnetization m_b here simply corresponds to $\rho_{\mathrm{gas}}^{\mathrm{coex}}$ in Fig, 8.1b, when the analogy between magnetization and density in the lattice gas interpretation of the Ising model is remembered. So m_s simply corresponds to ρ_s, and the result $m_s \propto t^{-1}$ then simply means that the average distance of the interface ℓ from the wall (in the partial wetting regime near the wetting transition) scales like t^{-1}. This is the same exponent as for $\xi_\perp$, cf. (8.19), $\nu_\perp = 1$. This information from (8.19)–(8.24) about the scaling theory of critical wetting can now be used as input for the appropriate extension of finite size scaling theory. Obviously, semi-infinite systems do not fit into the computer, so understanding the finite size effects that will be observed here is of crucial importance for a valid simulation study of critical wetting phenomena.

Anisotropic Finite Size Scaling

We recall from the discussion of Fig. 8.2 that for $T < T_w$ we will have a nonzero magnetization in the system, which can be positive (if the interface is bound to the lower wall) or negative (if it is bound to the upper wall), while for $T > T_c$, the magnetization for large enough L and M should be close to zero. Thus, it makes sense to simply discuss again the distribution of the total magnetization m, analogous to $P_L(s)$ in (8.14), but now we need to take into account that we have two inequivalent correlation lengths $\xi_\perp, \xi_{||}$ and two different linear dimensions L and M. It is natural to postulate that the order parameter scaling function $P_{L,M}(m)$ will depend on the ratios $L/\xi_\perp$ and $M/\xi_{||}$. Since $\xi_\perp^{\nu_{||}/\nu_\perp} \propto \xi_{||}$, we can use $(L/\xi_\perp)^{\nu_{||}/\nu_\perp}\xi_{||}/M = L^{\nu_{||}/\nu_\perp}/M$ as an alternative second argument of the scaling function: $L^{\nu_{||}/\nu_\perp}/M = c$ is a generalized aspect ratio of the system. The advantage of this transformation of the scaling function is that then the temperature dependence only enters via $\xi_{||}$. Thus the scaling ansatz that was postulated [8.56] is

$$P_{L,M}(m) = \xi_{||}^{\beta/\nu_{||}} \tilde{p}(L^{\nu_{||}/\nu_\perp}/M, M/\xi_{||}, m\xi_{||}^{\beta/\nu_{||}}) \quad . \tag{8.26}$$

It is important to realize that the exponent β that is introduced here is not the exponent β_s, considered in (8.25), just as the order parameter m used is the total

magnetization (per spin) of the thin film (or strip, in $d = 2$), rather than the surface excess magnetization of (8.23). Using again the normalization of $P_{L,M}(m)$,

$$\int_{-1}^{+1} P_{L,M}(m)\mathrm{d}m = 1 \tag{8.27}$$

we readily obtain the various moments

$$\langle |m|^k \rangle = \xi_{||}^{-k\beta/\nu}\, \tilde{m}_k(c, M/\xi_{||}) \quad k = 1, 2, \ldots \tag{8.28}$$

and hence the susceptibility for $T \leq T_\mathrm{w}$ becomes

$$k_\mathrm{B} T \chi' = L M \xi_{||}^{-2\beta/\nu_{||}}\, \tilde{\chi}(c, M/\xi_{||}) = c^{\nu_\perp/\nu_{||}}\, M^{1+\nu_\perp/\nu_{||}-2\beta/\nu_{||}}\, \tilde{\tilde{\chi}}(c, M/\xi_{||}) \tag{8.29}$$

where we have used the generalized aspect ratio c (which must be held fixed when M is varied) to eliminate L here; the scaling functions $\tilde{m}_k$, $\tilde{\chi}$, $\tilde{\tilde{\chi}}$ need not be specified explicitly here. From (8.29) we find that at the critical wetting transition in $d = 2$ we must have

$$k_\mathrm{B} T_\mathrm{w} \chi'_{T=T_\mathrm{w}} \propto M^{1+\nu_\perp/\nu_{||}-2\beta/\nu_{||}} = M^{3/2-2\beta/\nu_{||}} \quad . \tag{8.30}$$

On the other hand, we may consider for a system which is infinite in y-direction directly the finite size scaling of the surface susceptibility χ_s in (8.24),

$$\chi_\mathrm{s} = t^{-\gamma_\mathrm{s}} \tilde{\chi}(M/\xi_{||}) \propto \xi_{||}^2 \tilde{\chi}_s(M/\xi_{||}) \propto M^2 \,(T = T_\mathrm{w}) \quad . \tag{8.31}$$

Since the divergence of χ' simply comes from the divergence of χ_s, we can conclude

$$k_\mathrm{B} T_\mathrm{w} \chi' \mid_{t=0} = k_\mathrm{B} T_\mathrm{w} \chi_\mathrm{s} \mid_{t=0} / L = k_\mathrm{B} T_\mathrm{w} \chi_\mathrm{s} \mid_{t=0} / (M^{1/2} c^{-1/2}) \propto M^{3/2} \quad . \tag{8.32}$$

Comparing (8.30) and (8.32) we conclude that the exponent β introduced in (8.26) simply has the value $\beta = 0$; i.e. we have a critical phenomenon, but the order parameter in the thermodynamic limit disappears discontinuously.

Qualitatively this anomalous critical behavior is easily interpreted: when L is very large and $\ell \propto t^{-1}$ is still finite, the reduction of the average relative magnetization $\langle |m| \rangle / m_\mathrm{b}$ relative to unity is extremely small, $\langle |m| \rangle / m_\mathrm{b} \approx 1 - \ell/L$, m_b being the absolute value of the magnetization in the interior of the two coexisting domains in Fig. 8.2. Thus, in the limit $L \to \infty$ we have $\langle |m| \rangle = m_\mathrm{b}$ for all $t > 0$, while $\langle |m| \rangle = 0$ for $t < 0$. Note that for $T = T_\mathrm{w}$, $t = 0$, all moments $\langle |m|^k \rangle_{T_\mathrm{w}} = \tilde{m}_k(c, 0)$ take non-trivial values: at $T = T_\mathrm{w}$, the interface in the geometry of Fig. 8.2 is neither localized

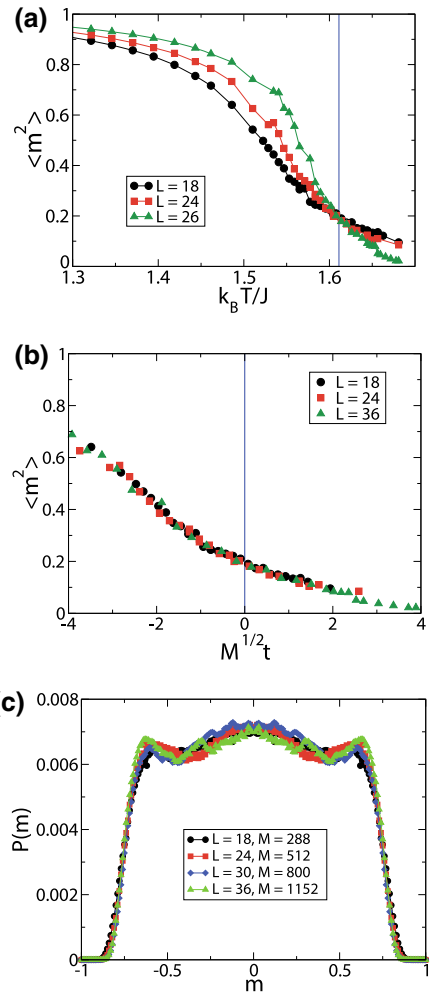


Fig. 8.3 **a** Plot of the squared magnetization $\langle m^2 \rangle$ versus temperature for the two-dimensional Ising model on the square lattice in a $M \times L$ geometry, Fig. 8.2 choosing $L^2/M = c = 9/8$ and $H_1/J = 0.70$, showing three choices of L, as indicated. The vertical straight line indicates the exactly known temperature T_w [8.57–8.59] of the critical wetting transition. **b** Scaling plot of the data for $\langle m^2 \rangle$ shown in (**a**) versus $M^{1/2}t$ to demonstrate that according to (8.28) $\langle m^2 \rangle$ is a function of $M/\xi_{||} \propto Mt^2$ only. **c** Distribution $P_{L,M}(m)$ versus m for the choice of parameters mentioned in (**a**), for $T = T_w$, to demonstrate the anomalous macroscopic fluctuations of the order parameter. From [8.56]

at one of the walls nor is it localized in the center of the film (or strip, respectively in $d = 2$ dimensions): the interface exhibits anomalous macroscopic fluctuations over the entire system. The values $\tilde{m}_k(c, 0)$ reflect the nontrivial distribution of these fluctuations, which depends on the generalized aspect ratio c.

A consequence of this behavior is that T_w can be found from searching for common intersection points of the moments $\langle |m| \rangle$, $\langle m^2 \rangle$ etc. plotted versus temperature, not only from intersections of the fourth-order cumulant $U_{L=1} - \langle m^4 \rangle / [3\langle m^2 \rangle^2]$, as is usually done. Figure 8.3 demonstrates this behavior for the $d = 2$ Ising model [8.56], where the exact solution for T_w is known (8.20). Of course, the value of this analysis presented in (8.26)–(8.32) is not just that one can reproduce exact results (8.19), (8.20), (8.22) and (8.25), but this methodology also readily works for systems

for which no exact solutions is available, such as the Blume-Capel model [8.72, 8.73] with surface fields. The Hamiltonian of this model is formulated in terms of spins S_i that can take three values $S_i = \pm 1$ and 0,

$$\mathcal{H} = -J \sum_{\langle i,j \rangle} S_i S_j + \mathcal{D} \sum_i S_i^2 - H \sum_i S_i - H_1 \sum_{i \in \text{row } 1} S_i - H_L \sum_{i \in \text{row } L} S_i \quad . \tag{8.33}$$

The first three terms on the righthand side of (8.33) represent the Blume-Capel model in the bulk: J is the exchange coupling between nearest neighbor spins on the lattice, H is the bulk magnetic field, and the parameter D controls the density of "vacancies" (i.e., sites i with $S_i = 0$). For $D \to -\infty$, vacancies are completely suppressed, and hence the model reduces to the standard Ising model as a limiting case. The surface fields H_1, H_L have already been introduced in the context of Fig. 8.2. In Fig. 8.4 we present numerical results for $\langle m^2 \rangle$ and $P_{L,M}(m)$ for the model: with the basic knowledge on Monte Carlo simulations for simple lattice models as exposed in detail in Chaps. 2 and 3 of this book, it should be a simple exercise for the reader to reproduce Figs. 8.3 and 8.4a, b on his own (at least for not too large lattice linear dimensions). From Fig. 8.4c it is evident that the finite size analysis tools described in this section do allow a convenient estimation of wetting transition temperatures for a broad class of models exhibiting critical wetting.

We end this subsection with a few comments: (i) just as β in (8.26)–(8.30) denotes an order parameter exponent for the total magnetization of the film, we also can introduce a susceptibility exponent γ for the total susceptibility (such as χ' in (8.30)–(8.31)). Since we have $\chi'_{T=T_w} \propto M^{\gamma/\nu_{||}}$, (8.32) implies $\gamma = 3$. This implies that the standard scaling relation $\gamma + \beta = \Delta_s = 3$ holds, as well as the hyperscaling relation extended [8.74] to anisotropic critical phenomena, $\nu_{||} + (d-1)\nu_\perp = 2\beta + \gamma = 3$. (ii) Other critical phenomena with anomalous interfacial fluctuations and an order parameter exponent $\beta = 0$ occur for the so-called "filling transitions", such as condensation of a fluid filling a wedge (e.g. [8.75, 8.76], or a corner (in $d = 2$) [8.77] or conical or pyramidal pores [8.78]). However, these problems are too specialized to merit a discussion here. (iii) The jumpwise disappearance of the order parameter ($\beta = 0$) found here is reminiscent of the behavior at bulk 1st order transitions, and indeed in the latter case we have also found common intersection points of the moments of the order parameter distribution (8.7). Also in this case, there did occur an anomalous macroscopic fluctuation, the system can jump between all coexisting states. However, there are well identifiable distinct states, the order parameter distribution in the thermodynamic limit simply is a sum of delta functions, unlike the case considered here {(8.26) for $L \to \infty$ is a nontrivial distribution $\tilde{p}'(c, mL^{\beta/\nu_{||}}) = \tilde{p}'(c, m)$, see Fig. 8.3c for an explicit example}. Thus, the common intersection points of the order parameter moments and cumulants for critical wetting cannot be simply predicted, unlike the case of first-order transitions. (iv) There exist also cases of anisotropic bulk critical phenomena, where two different critical exponents $\nu_{||}$, $\nu_\perp$ occur, but $\beta > 0$. A well-known example is the so-called "Lifshitz point" [8.79], that can be found in the so-called "axial next nearest neighbor Ising (ANNNI) model [8.80], for instance. Other anisotropic critical phenomena where

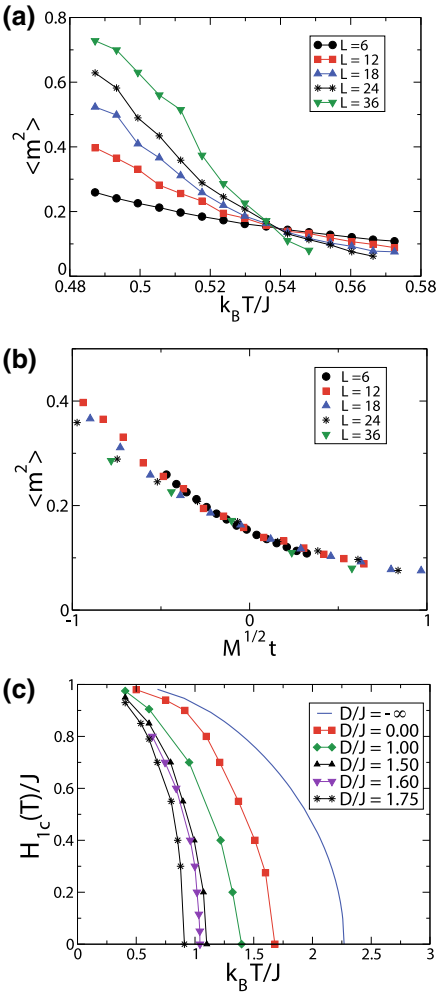


Fig. 8.4 **a** Plot of the squared magnetization $\langle m^2 \rangle$ versus temperature for the two-dimensional Blume-Capel model on the square lattice in a $M \times L$ geometry, choosing $c = 9/8$, $H_1/J = 0.85$, $\mathcal{D}/J = 1.75$, and 5 choices of L as indicated. From this data (and related ones for $\langle |m| \rangle$ and U_L) the critical wetting temperatures was estimated as $k_B T_w(H_1)/J = 0.538 \pm 0.004$. **b** Scaling plot of the data of (**a**) versus $M^{1/2}t$, to demonstrate that $\nu_{||} = 2$ works also for the Blume-Capel model, as expected from the universality principle of critical phenomena [8.22, 8.23]. **c** Plot of $H_{1c}(T)/J$, the inverse function to $k_B T_w(H_1)/J$, versus temperature for a range of values of D/J, as indicated. The result for $D/J = -\infty$ is (8.20). From [8.56]

(8.26) is applied occur in driven systems far from thermal equilibrium [8.81]. The "canonic" example is the driven lattice gas model,. where particles carry an electric charge and a field is applied creating a current in a chosen lattice direction (but Coulomb interactions between the particles are not considered) [8.82]. The transition to the ordered non-equilibrium steady state of this model can also be analyzed in terms of (8.26) [8.74, 8.83].

8.3 Interface and Surface Excess Free Energies and the Associated Finite Size Effects

We have already seen in Sect. 8.2.3, that interface and wall tensions play a crucial role for understanding wetting phenomena. Of course, these quantities play a crucial role also for the understanding of nucleation phenomena [8.85–8.88], i.e. the decay of a metastable phase by the formation of a (nanoscopic) droplet of the new phase, either via statistical fluctuations alone ("homogenous nucleation") or facilitated e.g. due to reduction of the free energy barrier if the droplet is attached to a wall, as in Fig. 8.1 ("heterogeneous nucleation"). The free energy barrier that needs to be overcome in a nucleation event is controlled by a competition of bulk and interfacial free energy contributions, and accurate estimation of the latter is crucial for the prediction of nucleation rates. This is a difficult task, even in the framework of the simple Ising/lattice gas model [8.88]. We also note that the interfacial and surface phenomena play a key role for many aspects of nanoparticles and their use in nanotechnology. Therefore, there has been a strongly increasing interest in recent years to study interfacial and surface excess free energies by simulation methods. In this section, we again focus on a few selected methodological aspects of this problem, which already are encountered in the study of the simple model systems emphasized in this book, the Ising/lattice gas system and the Lennard-Jones fluid.

Excess Free Energy of Interfaces Between Coexisting Phases

To introduce the subject, it again is useful to consider the simple Ising ferromagnet with nearest neighbor exchange interaction, which is one of the "workhorses" of this book, but now we specialize to a geometry of the lattice where one linear dimension (L_z) is distinctly larger than other(s), (L). We now focus on various choices of boundary conditions in the z-direction, while in the other direction(s) we take the standard periodic boundary conditions (PBC), see Chap. 2. Suppose we take free boundary conditions in z-direction but replace the missing rows (in $d = 2$) or planes (in $d = 3$) by rows (or planes) of spins fixed at $S_i = +1$ or $S_i = -1$, respectively (Fig. 8.5a). Then for temperatures T less than the critical temperature T_c, we expect that such competing boundary conditions enforce the coexistence of two domains, one with positive magnetization m_+ per spin (fluctuating around the value of the spontaneous magnetization $+m_{coex}$, symbolized by the double arrow pointing upward) and another domain with negative magnetization m_- per spin with m near $-m_{coex}$ (double arrow pointing downward). Note that at zero magnetic field both the size of the domains in z-direction, the values of m_+, m_-, and the local position of the interface can show fluctuations, as will be discussed below. When we compare this situation with systems with symmetric choices $(++)$ and $(--)$ of fixed spin boundaries, the situation will just differ by the presence of an interface in the case of the $(+-)$ boundary conditions. This consideration suggests to define the interfacial tension as excess free energy per unit length (in $d = 2$) or unit area (in $d = 3$) of the interface in terms of the appropriate combination of total free energies in the system, in d dimensions

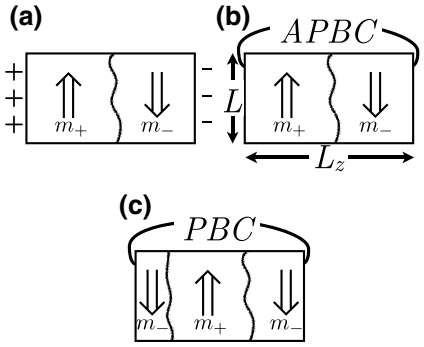


Fig. 8.5 Useful boundary conditions to study interfaces in Ising systems, using a simulation geometry of linear dimension(s) L parallel to the interface(s) and L_z in the direction perpendicular to the interface. In the parallel direction(s), periodic boundary conditions (PBC) are used throughout. Case (**a**) uses free surfaces in the layers $n = 1$ and $n = L_z$, while spins in the layer $n = 0$ are fixed at $S_i = +1$ (denoted by +), and $S_i = -1$ (denoted by −) in the layer $n = L_z + 1$ (implying a convention where the lattice spacing is taken to be unity). For $T < T_c$, two domains with positive magnetization (m_+) and negative magnetization (m_-) are then expected to coexist, separated by an interface (wavy line). Case (**b**) shows, as an alternative, the anti-periodic boundary condition (APBC), cf. text. With fully periodic boundary conditions (case **c**) one also can have coexistence between the two domains with opposite orientation of the magnetization, but then two interfaces must occur to satisfy the PBC. From Schmitz et al. [8.89]

$$\sigma(L, L_z) = L^{-(d-1)}\{F_{+-}(L, L_z) - [F_{++}(L, L_z) + F_{--}(L, L_z)]/2\} \quad . \tag{8.34}$$

Note that the factor 1/2 simply accounts for the fact that in the systems with symmetric choices of the fixed spins (++), (−−) we have these fixed spin boundaries twice, while both of them occur only once for the (+−) system. Of course, for the Ising model we would have the symmetry $F_{++} = F_{--}$; but the general formula, (8.34), has the merit, that it can easily be generalized to asymmetric cases, e.g. a standard Lennard-Jones fluid: then (++) stands for walls with short-range attraction to the particles, favoring the liquid phase of the fluid, while (−−) stand for walls with a short-range repulsion, favoring the gas; (8.34) is a "first principles" definition, it does not imply any assumptions on the local structure of the interfacial region. For the Ising model (or other symmetric systems, such as the symmetric binary (A,B) Lennard -Jones mixture, see e.g. [8.90]) an elegant alternative is the antiperiodic boundary condition (APBC), defined by [the index i of a lattice site stands for the indices $i_1, i_2, \ldots, i_d$ in the d directions of the hyper cubic lattices], see Fig. 8.5b

$$s(i_1, \ldots, i_{d-1}, i_{d\pm}L_z) = -S(i_1, \ldots, i_{d-1}, i_d) \tag{8.35}$$

i.e. the sign of the spin is reversed when the interaction crosses a boundary in z-direction. This APBC is fully equivalent to a system where the bonds in z-direction have been changed from $+J$ to $-J$ between one pair of rows (or planes, respectively). Then instead of (8.34) one uses

$$\sigma(L, L_z) = L^{-(d-1)}\{F_{\mathrm{APBC}}(L, L_z) - F_{\mathrm{PBC}}(L, L_z)\} \quad . \tag{8.36}$$

We have deliberately kept the arguments (L, L_z) for σ in (8.34) and (8.36) to emphasize that one needs to consider finite size effects. As we shall see, one source of finite size effects is the translational entropy of the interface in the z-direction: assuming an ensemble where the magnetic field H is fixed at $H = 0$ (so the linear dimensions of both positively magnetized and negatively magnetized domains along the z-directions in Fig. 8.5 on average will be equal) the interface position along the z-axis is not fixed, but will fluctuate. This translational degree of freedom creates an entropy $-k_{\mathrm{B}} T L^{-(d-1)} \ln(L_z/\ell_0)$ to σ, where the normalizing length ℓ_0 is often assumed to be the lattice spacing. While in the APBC case the interface can be anywhere along the z-axis, in the case of Fig. 8.5a the interface will not be close to one of the fixed spin boundaries. Thus one can expect that this entropic contribution will differ between the two cases Fig. 8.5a, b.

With PBC also in z-direction only even numbers of interfaces can occur, the relevant cases being no interfaces (the system being in states with uniformly positive or negative magnetization or 2 interfaces (as shown in Fig. 8.5c). Then

$$\sigma(L, L_z) = \frac{1}{2} L^{-(d-1)} \left\{ F^{(2)}_{\mathrm{PBC}}(L, L_z) - \left[F^{(0,+)}_{\mathrm{PBC}} + F^{(0,-)}_{\mathrm{PBC}} \right]/2 \right\} \quad . \tag{8.37}$$

Here $F^{(0,+)}_{\mathrm{PBC}}$ and $F^{(0,-)}_{\mathrm{PBC}}$ denote the free energies of systems with zero numbers of interfaces, with uniform positive or negative magnetization. Of course, in extremely elongated systems ($L_z \gg L$) one may encounter the need to consider systems with more than 2 interfaces; this problem is relevant for capillary condensation in long cylindrical pores [8.91] but out of consideration here.

Again we state that for an Ising system $F^{(0,+)}_{\mathrm{PBC}} = F^{(0,-)}_{\mathrm{PBC}} = F^{(0)}_{\mathrm{PBC}}$ because of spin reversal symmetry, and so the curly bracket in (8.37) then simplifies to $F^{(2)}_{\mathrm{PBC}}(L, L_z) - F^{(0)}_{\mathrm{PBC}}(L, L_z)$; but (8.37) as it stands is more readily generalized to any systems that lack a symmetry between the coexisting phases, such as a Lennard-Jones fluid.

Whatever geometry is chosen, the problem remains to compute the free energies of two (or three) systems of volume $L_z L^{d-1}$ and take the appropriate difference to isolate the interfacial term proportional to L^{d-1}. Obviously, this approach is promising only if free energies can be computed with extremely good accuracy. A general introduction into Monte Carlo methodologies to sample free energies and free energy differences has already been provided in Chap. 6 of this book. Since the free energy is not a direct output of the Metropolis algorithm, unlike the internal energy or other derivatives of the free energy, see Chap. 2, the approach is not completely straightforward. Nevertheless, the straightforward thermal integration method, where free energy differences are computed as integrals of the internal energy $E(\beta)$ considered as function of the inverse temperature $\beta = (k_{\mathrm{B}} T)^{-1}$, has been useful. Alternatively, one uses $F = E - TS$ and computes the entropy S, from the specific heat $C(T)$ as $S(T) = \int_0^T [C(T')/T'] \mathrm{d}T'$ [8.92]; such methods indeed have yielded useful results for the interfacial free energy of the Ising model (with interfaces perpendicu-

lar to the z-axis of the lattice) [8.93–8.95]. This thermodynamic integration method can also be extended to compute the interfacial free energy $\sigma(\vartheta)$ as function of an angle ϑ that the interface makes with xy plane [8.96]. In this case (in $d = 3$) one uses PBC in y-direction, APBC in z-direction, screw periodic boundary conditions (SPBC) in x-direction (this means a shift of PBC by N_ϑ rows in z-direction, such that $\tan\vartheta = N_\vartheta/L$). Due to the lattice structure, the interface tension of the Ising model is a function of both polar angles, $\sigma = \sigma(\vartheta, \varphi)$, but the full angular dependence is not yet known explicitly. The problem is complicated by the fact that Ising model interfaces in $d = 3$ undergo a roughening transition (which for the (001)-interface on the simple cubic lattice occurs for $k_B T_R/J \approx 2.454$ [8.97]) which is associated with a diverging correlation length (of interfacial height fluctuations relative to the average interface plane) [8.98, 8.99]. This diverging correlation length causes both strong and long lived fluctuations ("critical slowing down" [8.55]) and pronounced finite size effects.

Before we describe the methods that were used in more detail, we show selected results on $\sigma(\theta)$ in Fig. 8.6 [8.100–8.102]. For $T > T_R$ the angular dependence is rather weak, its maximal value (for $\theta = \pi/4$) is enhanced relative to unity by just a few percent, and for $T \to T_c$ this angular dependence disappears altogether (but there both the order parameter distinguishing the coexisting phases and the interfacial tension vanish, of course [8.22, 8.23]). For $T < T_R$, on the order hand, the anisotropy rises strongly (Fig. 8.6b) and reaches a maximum value of $\sqrt{2}$ (there are twice as many broken bonds for a (110) interface than for a (100) interface, but the area of the (110) interface is larger by a factor of $\sqrt{2}$). Obviously, for $T \leq 1$ one is close to this asymptotic value.

Neither $\sigma(0)$ nor the anisotropy ratio $\sigma(\pi/4)/\sigma(0)$ show the least sign of a singularity at $T = T_R$. This fact is theoretically expected [8.98, 8.99]: the roughening transition is neither first nor second order, but of "infinite order" (the radius of convergence of a power series in $(T - T_R)$ would be zero). Of course, such a weak singularity clearly is not detectable by Monte Carlo simulations! However, there is a singularity that is better detectable in the angular dependence: for $T \geq T_R$ $\sigma(\theta)$has a quadratic expansion

$$\sigma(\theta) = \sigma(0)(1 + C\theta^2 + \cdots) \quad , T \geq T_R \tag{8.38}$$

while for $T < T_R$ there occurs a kink singularity

$$\sigma(\theta) = \sigma(0)(1 + C'|\theta| + \cdots) \quad , \quad T < T_R \quad . \tag{8.39}$$

The quantity $\tilde{\sigma}$,

$$\tilde{\sigma} = \sigma(0) + \frac{\mathrm{d}^2\sigma(\theta)}{\mathrm{d}\theta^2}\,|_{\theta=0} = \sigma(0)(1 + 2C) \tag{8.40}$$

is called the "interfacial stiffness" [8.103] and this quantity plays a central role in capillary wave theory [8.1–8.4, 8.49, 8.104]. Theory predicts that in $d = 3$ dimensions

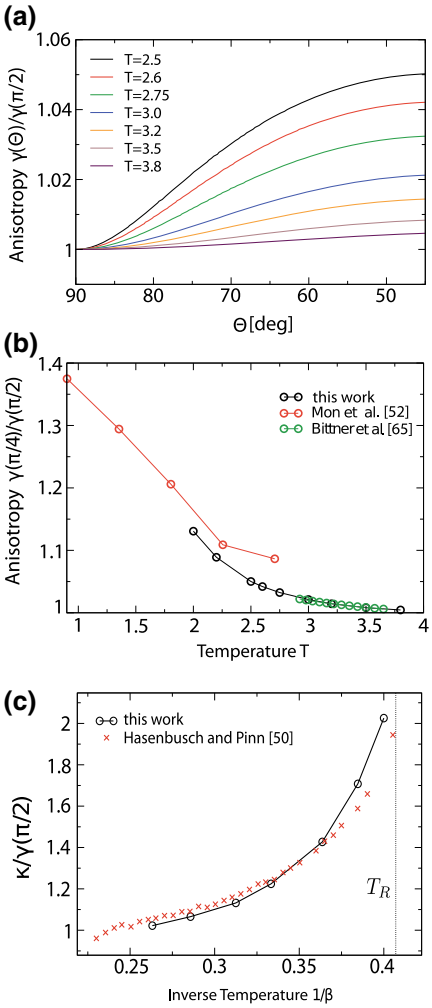


Fig. 8.6 **a** Dependence of the interface tension $\gamma(\theta)$ normalized by $\gamma(\pi/2)$ plotted versus the inclination angle of the interface relative to a lattice plane (note $\sigma(\theta) = \sigma(\pi/2 - \theta)$ by the symmetry of the simple cubic lattice). Several temperatures T (in units of J/k_B) are included, from $T = 2.5$ (top) to $T = 3.8$ (bottom). In the limit $T \to T_c$ ($\cong 4.5115$) [8.100]) $\sigma(\theta)$ becomes constant, independent of θ, but also vanishes according to the power law $\sigma(0) = \hat{\sigma}(1 - T/T_c)^{(d-1)\nu}$, where $\hat{\sigma}$ is a critical amplitude. **b** Maximum anisotropy $\sigma(\pi/4)/\sigma(0)$ plotted versus temperature. Upper curve is from Mon et al. [8.96] and overestimates the correct result due to the use of too small lattice linear dimensions for T near $T_R \approx 2.454$. The lower curve is based on lattices with $L_x \times L_y \times L_z = 184 \times 504 \times 504$ lattice sites. For $T \geq 2.9$ data due to Bittner et al. [8.101] obtained with the probability distribution method are included, and agree perfectly with the data from thermodynamic integration. **c** Ratio of interfacial stiffness κ and interface tension $\sigma(0)$ plotted versus inverse temperature. Circles are deduced from the data on $\sigma(\theta)$ via (8.40), while crosses are due to a capillary wave analysis of Hasenbusch and Pinn [8.94]. Straight line indicates the roughening transition. From Block et al. [8.102]

for $T \to T_R$ $\tilde{\sigma}(T)J/k_BT = (\pi/2)\{1 - C''[(T - T_R)/T_c]^{1/2} + \cdots\}$, while $\tilde{\sigma}(T < T_R) = \infty$. The data shown in Fig. 8.6c are compatible with such a behavior, but they are certainly no strong evidence yet.

The singular behavior of (8.39) means that an interface for $T < T_R$ at small $|\theta|$ looks like a regular staircase of broad terraces (of width proportional to $1/|\theta|$) separated by steps having the height of the lattice spacing a, apart from small local fluctuations. It is useful then to consider the step free energy $f_s(T)$, since one can show that $f_s(T \to T_R) \to 0$. This vanishing of $f_s(T)$ can also be studied by Monte Carlo methods [8.96, 8.102] but again a finite size rounding of this singular behavior must be considered. The vanishing of $f_s(T)$ means that there is a proliferation of surface step formation, and due to the superposition of many steps with irregular distances between them the distance of the local interface position from the considered reference plane in a particular interface configuration can be very large. Fig. 8.2 already gave a schematic picture of such a coarse-grained rough interface (confined between parallel walls) in $d = 2$ dimensions. It must be stressed, however, that in $d = 2$ we have $T_R = 0$, hence then the regime described by (8.39) does not exist. But nevertheless the square lattice Ising ferromagnet is a very useful "testbed" for Monte Carlo methods to study interfacial phenomena, see e.g [8.38, 8.46, 8.48, 8.88, 8.91, 8.105], since both $\sigma(\theta, T)$ and $\sigma(T)$ are known exactly [8.103, 8.106, 8.107]. In particular, $\tilde{\sigma}(T) = k_BT \sinh\{\sigma(0, T)/k_BT\}$ implies that $\tilde{\sigma}(T \to 0) \to \infty$ since $\sigma(0, T) = 2J$, while $\tilde{\sigma}(T) \approx \sigma(0, T)$ near T_c where $\sigma(0, T) \ll k_BT$. The fact that $\tilde{\sigma}(T)$ reduces to $\sigma(0, T)$, or more generally that the angular dependence of $\sigma(\theta, T)$ vanishes near T_c, also implies that droplet shapes become circular (in $d = 2$) or spherical (in $d = 3$), but at general temperatures droplet shapes are nontrivial [8.107–8.109]. While this problem has an exact solution in $d = 2$, large scale Monte Carlo simulations are needed to study it in $d = 3$ [8.109].

While the study of the roughening transition [8.98, 8.99] has a long history (e.g. [8.110, 8.111–8.113]), and also more complicated models closer to real materials were occasionally studied (e.g. [8.114, 8.115]), addressing the roughening of real crystal surfaces and interfaces, there is still need for careful large scale Monte Carlo studies of this problem.

Capillary Waves

The interfacial stiffness $\tilde{\sigma}(T)$ is a key concept for describing interfacial fluctuations. As anticipated already in Fig. 8.2, one wishes to disentangle bulk and interfacial fluctuations by a suitable coarse-graining. In the vicinity of the critical point of the Ising model (or of a fluid undergoing a vapor to liquid transition), the characteristic length scale of order parameter fluctuations in the bulk is the correlation length ξ_b. A commonly used idea then is to apply a block analysis method based on a block linear dimension $B > \xi_b$ [8.116] [e.g., if the interface is oriented perpendicular to the z-axis, one divides the xy-planes normal to the z-axis into n_B^2 squares of linear dimension B each ($L = n_B$). For each system configuration that is analyzed one then considers rectangular columns with a volume $B \times B \times \Delta z$, with $\Delta z = z_+ - z_-$, and the values z_+, z_-, are chosen such that the local interface position z_G is in between z_- and z_+ for all blocks; here z_G is defined as Gibbs dividing surface [8.49] in each

block (see [8.117] for more details and explicit examples). In this way a local height variable $h(x, y) = z_G$ is obtained. Of course, the "snapshots" of those coarse-grained interface configuration depend very much on the "resolution of the microscope" with which this coarse-grained interface is viewed [8.117, 8.118]. For small B the picture is rather rugged; but always fluctuations of the local height $h(x, y)$ are present, with wavelengths over the full range $B \leq \lambda \leq L$.

The idea then is that for $B > \xi_b$ "overhangs" of the local interface position no longer occur, and the interface is locally sufficiently flat so that the energy cost of the interfacial fluctuations can be described by the capillary wave Hamiltonian [8.49, 8.104], in $d = 3$

$$\mathcal{H}_{cw} = \frac{1}{2}\tilde{\sigma} \int_0^L \mathrm{d}x \int_0^L \mathrm{d}y (\nabla h(x, y))^2 \quad . \tag{8.41}$$

For off-lattice fluids or fluid binary mixtures there would be no distinction between $\tilde{\sigma}$ and the interface tension σ; however, for lattice systems such as the Ising model (as well as for the case of liquid-solid interface [8.119–8.124]) it is $\tilde{\sigma}$ which appears as an energy parameter in (8.41).

Equation (8.41) can be brought to diagonal form by introducing Fourier transforms, $h(x, y) = \sum_{\vec{q}} h(\vec{q}) \exp[i(q_x x + q_y y)]$ to find

$$\mathcal{H}_{cw} = \frac{1}{2}\tilde{\sigma} \sum_{q_x = 2\pi/L}^{2\pi/B} \sum_{q_y = 2\pi/L}^{2\pi/B} q^2 |h(\vec{q})|^2 \quad . \tag{8.42}$$

For such problems the equipartition theorem yields $\frac{1}{2}\tilde{\sigma} q^2 \langle |h(\vec{q})|^2 \rangle = \frac{k_B T}{2}$, and hence the capillary wave spectrum is described by

$$\langle |h(\vec{q})|^2 \rangle = k_B T / (\tilde{\sigma} q^2) \quad . \tag{8.43}$$

Numerical tests of this relation (where $\tilde{\sigma} = \sigma$ was found independently [8.117, 8.118]) have verified (8.43) convincingly. Actually, for liquid-solid interfaces one broadly relies on the validity of this description, and uses (8.43) to estimate $\tilde{\sigma}$ by studying $\langle |h(\vec{q})|^2 \rangle$ in the simulations. Note that quantities such as $\langle |h(\vec{q})|^2 \rangle$ can be straightforwardly sampled by the Metropolis algorithm, although the slowness of the interfacial relaxation on large length scales is a problem ((8.43) does imply that a rough interface is in a sense a critical object, there is no finite correlation length $\xi_{||}$ parallel to the interface that would limit the long-wavelength fluctuations).

A consequence of (8.43) also is that the mean-square width s_L^2 of the coarse-grained interface is size dependent. With $\langle h(x, y) \rangle = 0$ one concludes

$$s_L^2 = \langle |h(x, y))|^2 \rangle = \sum_{\vec{q}} \langle |h(q)| \rangle^2 = \frac{k_B T}{\tilde{\sigma}} \sum_{\vec{q}} q^{-2} \quad . \tag{8.44}$$

Converting the sums over q_x, q_y in (8.44), cf. (8.42), into an integral yields

$$s_L^2 = \frac{k_B T}{2\pi\tilde{\sigma}} \int_{2\pi/L}^{2\pi/B} \mathrm{d}q/q = \frac{k_B T}{2\pi\kappa} \ln(L/B) \quad . \tag{8.45}$$

Now one may argue that due to the interfacial structure on length scales smaller than B one should allow for an "intrinsic width" ω_0 of the interface on the scale B. This intrinsic width is commonly included in terms of a convolution approximation with the Gaussian distribution, $P_L(h) = (2\pi s_L^2)^{-1/2} \exp(-h^2/2s^2)$ [8.125], to find for the mean-square width of the interface on the length scale L

$$w_L^2 = w_0^2 - \frac{k_B T}{4\tilde{\sigma}} \ln B + \frac{k_B T}{4\tilde{\sigma}} \ln L \quad . \tag{8.46}$$

This logarithmic variation of w_L^2 with L has been extensively studied in simulations (e.g [8.94, 8.113, 8.115, 8.117, 8.123, 8.126]) and was also used to estimate $\sigma/k_B T$. However, (8.46) obviously does not allow to identify w_0 unambiguously, since it is found to depend on the (somewhat arbitrary) parameter B [8.118]. On scales $B < \xi$, the distinction between bulk and interface fluctuations is not really uniquely possible [8.127], and extending the capillary wave model to describe this regime is till an active area of research [8.128, 8.129]. Note also that in $d = 2$ a treatment analogous to (8.41)–(8.44) is possible but yields

$$s_L^2 = \frac{k_B T}{2\pi\tilde{\sigma}} a \int_{2\pi/L}^{2\pi/B} \mathrm{d}q/q^2 = \frac{k_B T}{\tilde{\sigma}} La \quad . \tag{8.47}$$

Thus, in $d = 2$ interfacial fluctuations are even stronger than in $d = 3$. These strong interfacial fluctuations are also one of the sources of the finite size effects hampering the estimation of interfacial tensions, that we will discuss now.

Logarithmic Finite Size Effects on Interfacial Free Energies

As an example for the finite size effects that are observed, we start with the method to extract the interfacial free energy from the order parameter distribution $P_{L,L_z}(\rho)$, considering for simplicity the $d = 2$ lattice gas model at $\mu = \mu_{\mathrm{coex}}$ (i.e., in "magnetic language" used in Sects. 8.2 and 8.3 this corresponds to bulk magnetic field $H = 0$ and the peaks of P_{L,L_z} at $\rho = \rho_{\mathrm{v}} = (1 - s_\infty)/2$ and $\rho = \rho_\ell = (1 + s_\infty)/2$ correspond to the states with positive and negative spontaneous magnetization $\pm s_\infty$, of course). For densities ρ near the density ρ_{d} of the rectilinear diameter, $\rho_{\mathrm{d}} = (\rho_{\mathrm{v}} + \rho_\ell)/2$, there occurs a flat (almost strictly horizontal) minimum (see Fig. 8.7a). This minimum is due to a state as shown in Fig. 8.5c: two domains of opposite magnetization (or densities ρ_{v} and ρ_ℓ, respectively) coexist in the system, separated by two domain walls, which are oriented normal to the z-axis (if PBC in all lattice directions are used). Then the difference in $P_{L,L_z}(\rho_{\mathrm{v}})$ [or $P_{L,L_z}(\rho_\ell)$] and $P_{L,L_z}(\rho_{\min})$ just is due to

the presence of the two domain walls; thus [8.105]

$$P_{L,L_z}(\rho_{\min})/P_{L,L_z}(\rho_v) \propto \exp[-2\sigma(L;L_z)L^{d-1}/k_BT] \quad , \tag{8.48}$$

and hence from a sampling of $P_{L,L_z}(\rho)$ one can estimate $\sigma(L, L_z)$,

$$\sigma(L, L_z) = (2L^{d-1})^{-1} \ln[P_{L,L_z}(\rho_v)/P_{L,L_z}(\rho_{\min})] \quad . \tag{8.49}$$

Note that using in the x-direction screw periodic boundary conditions (SPBC) as defined in Sect. 8.4.1 one can obtain the interface free energy of interfaces tilted by a chosen angle ϑ, as done in [8.101] and included in Fig. 8.6b. Equation 8.49 has also been used for off-lattice systems such as the Lennard-Jones fluid (e.g. [8.130]) and other models [8.131–8.133]. Techniques to accurately sample probability distributions, that vary over many orders of magnitude have already been described in Chap. 6 of this book.

Figure 8.7a shows, however, that (at least in $d = 2$) there does occur a significant dependence on L_z. While in the case of APBC in the grand canonical ensemble this dependence just reflects the translational entropy of the interface, which can be located anywhere in the system, giving rise to a term $S = -(k_BT/L)\ln(L_z)$, Fig. 8.7b, the situation is more subtle in case of Fig. 8.7a. Already in the case of APBC in the canonical ensemble, the interface (on average) is fixed at $z = L_z/2$, to maintain that the density remains at $\rho = 1/2$. However, Fig. 8.7b reveals a term $S = -\frac{1}{2}(k_BT/L)\ln L_z + \text{const}$. This comes about by the "domain breathing effect" [8.89]. The "magnetizations" m_+, m_- in Fig. 8.6b can fluctuate, even when the average magnetization $\langle m\rangle = 0$ is fixed, by corresponding fluctuations of the interface position around its average position. These fluctuations imply the excess entropy quoted above. Now in the case of PBC (and the canonical ensemble, to ensure $\langle m\rangle = 0$) the domain with positive magnetization in Fig. 8.6c can be arbitrarily translated in the simulation box, yielding the above translational entropy, but also the distance between the walls can fluctuate, due to the above domain breathing effect. This yields altogether an entropy $-(3k_BT/2L)\ln L_z$, but since we have two interfaces in Fig. 8.6c, the contribution per interface then is $-(3k_BT/4L)\ln L_z$. This consideration interprets the three slopes 1/2, 3/4 and 1 demonstrated in Fig. 8.7b.

Understanding the L-dependence of $\sigma(L, L_z)$ even is more subtle. One can show [8.134] that in $d = 2$ capillary waves cause a correction of order $(k_BT/2L)\ln L$, but the domain breathing effect also contributes a term of order $L^{-(d-1)}\ln L$ [8.89]. A detailed discussion shows that (with the notation $\gamma_{L,L_z} \equiv \sigma(L, L_z)/k_BT$) [8.89]

$$\tilde{\gamma}_{L,L_z} = \gamma_{L,L_z} + \frac{x_\perp \ln L_z - x_{||} \ln L}{L^{d-1}} = \gamma_\infty + \frac{C}{L^{d-1}} \quad . \tag{8.50}$$

The constants $x_{||}$, $x_\perp$ are universal {$x_\perp$=1(APBC (gc)), 1/2 (APBC(c)) or 3/4 (PBC(c)), both in $d = 2$ and $d = 3$; $x_{||} = 1/2$ (APBC (gc), 1 (APBC (c) and 3/4 (PBC (c)), in $d = 2$, while $x_{||} = 0$, 1 1/2, in these three cases but $d = 3$, respectively}. The constant C is non-universal. But Fig. 8.7c shows that (8.50) indeed

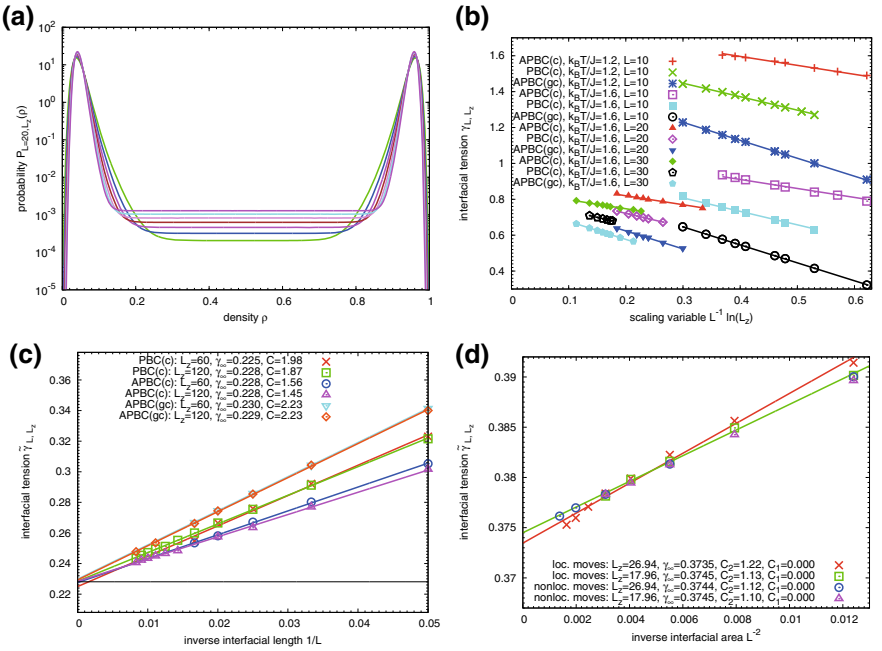


Fig. 8.7 **a** Density distribution $P_{L,L_z}(\rho)$ versus ρ for the two-dimensional Ising/lattice gas model at the temperature $k_BT/J = 2.0$ and the choice $L = 20$ for several choices of L_z : $L_z = 30, 40, 50, 60, 70, 80, 90$, and 100 (from bottom to top at $\rho = 0.5$) **b** Interfacial tension $\sigma(L, L_z)$ for the $d = 2$ Ising model at fixed L versus $L^{-1}\ln(L_z)$ for several choices of k_BT/J and L, as indicated in the key, and different combinations of boundary conditions (APBC, PBC) and statistical ensembles of the lattice gas: grand canonical (gc) and canonical (c); in the latter case, $\rho = 0.5$ is kept fixed. The straight lines shown have the slopes $x_\perp = 1/2$, 3/4 and 1, discussed in the text (**c**) Reduced interfacial tension $\tilde{\gamma}(L, L_z)$ plotted versus $1/L$ for the $d = 2$ Ising model for two choices of $L_z (L_z = 60, 120)$ and fitted to the form $\tilde{\gamma} = \gamma_\infty + C/L$, at $k_BT/J = 2.0$. The fitted results and choices of boundary conditions and ensembles (cf. part **b**) are quoted in the key. Horizontal straight line indicates the exact result [8.106] $\gamma_\infty = 0.228$. **d** Reduced interfacial tension $\tilde{\gamma}(L, L_z)$ of the $d = 3$ Lennard-Jones fluid at $k_BT/\varepsilon = 0.78$ plotted versus L^{-2}. Lengths are measured here in units of the Lennard-Jones diameter σ (and the potential is cut and shifted to zero at $r = 2.56\sigma$). Two choices of L_z are included, and the results for γ_∞ and $C = C_2$ are quoted in the key (requesting that there is no correction C_1/L here). Both local and nonlocal moves of the particles in the simulation box were used. From Schmitz et al. [8.89]

provides a good fit to the observed data, and converges to the exactly known result [8.106] for all three choices of boundary conditions. It is obvious from Fig. 8.7c that the data for $\tilde{\gamma}_{L,L_z}$ (and those for γ_{L,L_z} as well) do depend on the choice of boundary conditions and ensemble, and finite size effects in $d = 2$ are quite pronounced, Fortunately, due to the denominator L^2 (rather than L) in $d = 3$ the finite size effects typically are only of the order of a few percent. Figure 8.7d gives an example for the LJ fluid (due to the lack of symmetry between the coexisting phases then only the case of PBC(c) is possible).

At this point we emphasize that all the data shown in Fig. 8.7b–d were obtained with a rather new method to estimate the interfacial tension $\sigma(L, L_z)$, namely the so-called "ensemble switch" method [8.89, 8.135]. In this method one constructs the desired free energy difference between a system without interfaces and a system with one (or two) interfaces by defining a mixed Hamiltonian,

$$\mathcal{H} = \kappa \mathcal{H}_1 + (1 - \kappa)\mathcal{H}_0 \,, \quad 0 \le \kappa < 1 \quad . \tag{8.51}$$

Here $\mathcal{H}_1$ is the desired system with two interfaces, while $\mathcal{H}_0$ is a system which contains the two coexisting phases in separate boxes (of linear dimension $L_z/2$ in z-direction) each, with periodic boundary conditions. Figure 8.8 illustrates this choice for a Ising model with PBC. The two systems described by $\mathcal{H}_1$ and $\mathcal{H}_0$ hence have exactly the same total volume and contain exactly the same number of particles (Ising spins in the case of Fig. 8.8). Thus, although $\mathcal{H}(\kappa)$ cannot be realized in a laboratory experiment, from the point of view of statistical mechanics (8.51) is a perfectly well-defined Hamiltonian suitable for the application of Monte Carlo methods, irrespective of the value of the parameter κ. The systems for $\kappa = 0$ and $\kappa = 1$ only differ by the choice of boundary conditions, and hence

$$\sigma(L, L_z) = [F(\kappa = 1) - F(\kappa = 0)]/(2L^{d-1}) \quad . \tag{8.52}$$

By varying κ one hence can smoothly interpolate between the system with and without interfaces. In practice one chooses a set of n_κ discrete values κ; and computes free energy differences $\Delta F = F(\kappa_i) - F(\kappa_{i-1})$ by (successive) umbrella sampling [8.136], to find $\Delta F = \sum_{i=1}^{n_\kappa} \Delta F_i = F(\kappa = 1) - F(\kappa = 0)$. Here κ_i is taken as an additional variable in the Monte Carlo sampling, allowing for moves κ_i to κ_{i+1} or vice versa, to sample the probability $P(\kappa_i)$ relative to $P(\kappa_{i+1})$. Then $\Delta F = k_B T \ln\{P(\kappa_{i-1})/P(\kappa_i)\}$. Since the n_κ "windows" $\{\kappa_{i-1}, \kappa_i\}$ can be sampled independently of each other simultaneously, this algorithm is suited for very efficient parallelization.

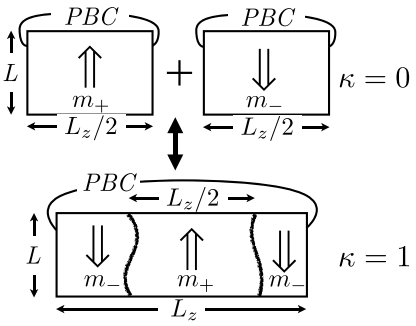


Fig. 8.8 **a** Schematic illustration of the ensemble switch method, for the case corresponding to Fig. 8.5c. From Schmitz et al. [8.89]

In a standard numerical integration, one would choose windows of equal size, $\kappa_i = i/n_\kappa$. However, it is better to choose the distribution of points κ_i; more closely spaced near $\kappa = 0$ and $\kappa = 1$, e.g. according to [8.135, 8.137]

$$\kappa_i = \sin^2(i\pi/2n_\kappa), i = 0, 1, \ldots, n_\kappa \quad . \tag{8.53}$$

We emphasize that this algorithm can also be used for off-lattice models, to study the interface tension of liquid-vapor interfaces [8.89, 8.136, 8.137] and liquid-solid interfaces.

For liquid-vapor interfaces (or for interfaces between coexisting phases in liquid mixtures, respectively) the standard approach to estimate the interfacial free energy has been based on the anisotropy of the pressure tensor $P_{\alpha\beta}(z)$ across the interface [8.49]

$$\sigma(L, L_z) = \frac{1}{2}\int_0^{L_z} \{p_{zz}(z) - \frac{1}{2}\Big[p_{xx}(z) + p_{yy}(z)\Big]\}\mathrm{d}z \quad . \tag{8.54}$$

Here again it was assumed that PBC are used throughout, i.e. two interfaces are present, and the canonical ensemble ensures that the density in the system is constant. The linear dimensions L, L_z are chosen such that the two interfaces are essentially non-interacting. Far from the interfaces, $p_{xx}(z) = p_{yy}(z) = p_{zz}(z)$ just is the bulk pressure in the fluid, hence there the integrand of (8.54) vanishes (in practice, of course, there is a problem to distinguish statistical fluctuations of the integrand from small but systematic effects in the wings of the interfacial profile) Although this traditional method [8.138] has been widely used, particularly in the context of Molecular Dynamics simulations where the forces that are needed to compute the pressure tensor from the Virial formula anyway are calculated to numerically integrate Newton's equation of motion of the particles, finite size effects hampering (8.54) have been considered only rarely [8.139]. A further drawback of (8.54) is that it cannot be used in cases where the pair potential between the particles is singular, e.g. for hard spheres, the AO model of colloid polymer mixtures [8.35, 8.36], etc. Alternative methods for computing the pressure tensor in such cases have been devised [8.140] but are cumbersome to apply.

Returning to Ising models, we note that many alternative methods to obtain the interfacial tension were discussed in the literature, but lack of space prevents us from describing all of them. We only mention one method [8.141, 8.142] based on Jarzyinski's nonequilibrium work theorem [8.143]

$$\left\langle \exp\Big[-\frac{W(t_\mathrm{i}, t_F)}{k_\mathrm{B}T}\Big]\right\rangle = \exp\Big(\frac{\Delta F}{k_\mathrm{B}T}\Big), \tag{8.55}$$

where a dynamic nonequilibrium process is considered that transforms the state of a system from an initial state (i) at time t_i to a final state (f) at time t_f, $W(t_\mathrm{i}, t_\mathrm{f})$ denotes the total work done on the system during this transformation, and $\Delta F = F_\mathrm{f} - F_\mathrm{i}$ is the desired free energy difference.

To apply (8.55) to the Ising model interfacial free energy, we consider as initial state a system with PBC throughout, which does not have any interfaces, and as final state we have a system where between two planes (or rows in $d = 2$, respectively) the exchange couplings are $-J$ instead of $+J$ (recall that this situation is equivalent to the choice of uniform interactions with APBC in z-direction). The nonequilibrium process that one considers is a gradual transformation of these exchanged couplings between the considered two special planes (rows) from $+J$ to $-J$, i.e.

$$J(t_\mathrm{n})/|J| = 1 - 2n/\mathcal{N}, \quad t_\mathrm{n} = t_\mathrm{i} + n(t_\mathrm{f} - t_\mathrm{i})/\mathcal{N} \quad , \tag{8.56}$$

where $\mathcal{N}$ is the number of intermediate systems between the initial and final system. The work $\delta W(t)$ done in a time interval $\delta t = (t_\mathrm{f} - t_\mathrm{i})/\mathcal{N}$ simply is related to the energy change $\delta\mathcal{H}$ produced by (8.56) at this particular instant t of this nonequilibrium process, and $W(t_\mathrm{f} - t_\mathrm{i}) = \int_{t_\mathrm{i}}^{t_\mathrm{f}} \delta W(t)\mathrm{d}t$. Of course, the sampling of $\exp(-W/k_\mathrm{B}T)$ in (8.55) still requires a major effort, and it was found [8.142] that $\mathcal{N}$ has to be large as well, $\mathcal{N} \approx 10^5$, to yield accurate results. In this algorithm, it is rather natural to run this process in the backward direction in time as well, and check that both directions yield the same result. Obviously, for the case of Ising systems with APBC this method is closely related to the above ensemble switch method, since (8.56) just creates "mixed Hamiltonians" $\mathcal{H}(\kappa)$ as used in (8.51).

Many special methods gradually creating interfaces have also been developed for the case of interfaces between crystals and liquids (see e.g. [8.144, 8.145] and references therein). However, this topic is beyond our scope here.

8.4 Wall Excess Free Energies

For the understanding of wetting phenomena (Sect. 8.2.3), heterogeneous nucleation at walls [8.42, 8.43–8.47], etc, the interfacial tension between the appropriate coexisting phases is one ingredient, but a second ingredient is the difference in the appropriate surface free energies due to the wall (cf. Fig. 8.1 and (8.18), for instance).

In the Ising model, the spin reversal symmetry can again be exploited in the context of thermodynamic integration methods to find directly the contact angle (but one needs to be careful due to the anisotropy of the interfacial free energy, discussed in Sect. 8.3). Block et al. [8.102] proposed to study this problem using Ising systems with linear dimensions L_x, L_y and L_z, where on the bottom $L_z \times L_y$ surface a surface field $-|H_1|$ acts, while on the top $L_z \times L_y$ surface a surface field $+|H_1|$ acts. In z-direction PBC are used. If $H_1 = 0$ and APBC are used in y-direction, an interface perpendicular to the y-direction is stabilized. For $|H_1| > 0$ this interface is inclined by an angle θ relative to the yz-plane (Fig. 8.9a). However, it then is imperative to use a so-called "generalized antiperiodic boundary condition" (GAPBC) [8.102]

$$S(x, y, z) = -S(L_x - x, y \pm L_y, z) \tag{8.57}$$

where x, y, z are the Cartesian coordinates of the spin S_i; elsewhere labelled simply by the index i {cf. (8.33)}. Note that for the choice of surface fields $H_n = -H_1$ the surface excess free energies $f_1^{(+)}(T, H_1)$, $f_1^{(-)}(T, H_1)$ of the system with positive (+) or negative (−) magnetization in the bulk due to the bottom wall (layer index $k = 1$) are related by symmetry to the corresponding properties $f_n^{(+)}(T, H_n)$, $f_n^{(-)}(T, H_n)$ of the top wall,

$$f_1^{(+)}(T, H_1) = f_n^{(-)}(T, H_n) \quad , \quad f_1^{(-)}(T, H_1) = f_n^{(+)}(T, H_n) \tag{8.58}$$

and hence we find for the local magnetization via the derivatives $m_1 = -(\partial f_1/\partial H_1)_T$, $m_n = -(\partial f_n/\partial H_n)_T$ that

$$m_1^{(+)}(H_1) = -m_n^{(-)}(H_n) \quad , \quad m_1^{(-)}(H_1) = -m_n^{(+)}(H_n) \quad . \tag{8.59}$$

This is a special case of the more general symmetry $m^{(+)}(x, H_1, H_n = -H_1) = -m^{(-)}(L_x - x, H_1, H_n = -H_1)$ resulting from (8.57), cf. Fig. 8.9b, where for simplicity x was treated as if it is a continuous variable (rather than being defined only on lattice sites). Thus, the effect of GAPBC is that the coexisting domains are of quasi-infinite size, their profiles in x-direction far from the interface are those of bulk domains, exhibiting the proper symmetries, (8.58) and (8.59). Since the free energy difference $\sigma_{wg} - \sigma_{w\ell}$ in Young's equation, (8.18), in the present notation is nothing but $\Delta f_1 = f_1^{(+)}(T, H_1) - f_1^{(-)}(T, -H_1)$, (8.58) and (8.59) imply that

$$\Delta f_1 = -\int_0^{H_1} dH_1'[m_1(H_1') + m_n(H_1')] \tag{8.60}$$

but as we shall see, Young's equation (that would take the form $\sigma(L, L_z)\cos\theta = \Delta f_1$ here) needs modification due to the anisotropy of the interfacial tension.

We now discuss the free energy of the Ising model for the geometry of Fig. 8.9a.

We subtract from the free energy of the system the free energy of a system with the same linear dimensions L_x, L_z but with PBC throughout, in a state with uniformly positive or negative spontaneous magnetization. The result of this subtraction will contain excess terms due to the walls and the inclined interface (and the two contact lines where the interface meets the wall). The interfacial area is $\tilde{L}L_z$ with $\tilde{L} = L_x/\sin\theta$. Making use of the symmetries noted in (8.58) this total interfacial free energy contribution can be cast in the form [8.102].

$$\begin{aligned} F_{\text{int}} &= L_x L_y \sigma(\theta)/\sin\theta + L_z(L_y + L_x/\tan\theta)\gamma^{(+)}(|H_1|) \\ &+ L_z(L_y - L_x/\tan\theta)\gamma^{(+)}(-|H_1|) + 2L_z\tau \end{aligned} \tag{8.61}$$

where the line tension τ has been introduced. Note that (8.61) simply reflects the geometry of Fig. 8.9a, and we have taken advantage of the special symmetry occurring

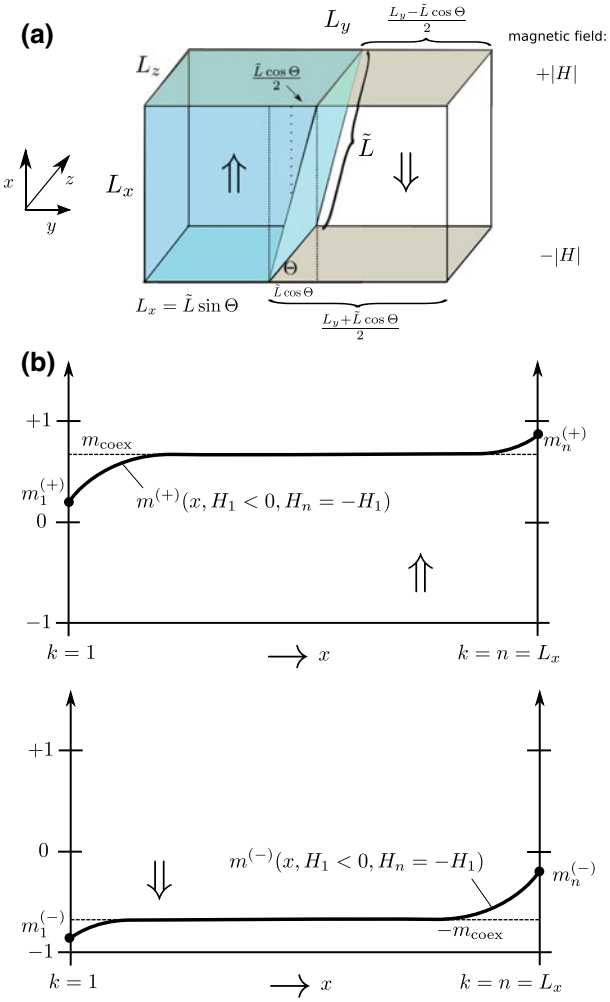


Fig. 8.9 **a** Simulation geometry for a simple cubic lattice with box linear dimensions L_x, L_y, L_z in the x, y, z-directions. While PBC are used in the z-direction, at the two free surfaces of linear dimensions L_y, L_z boundary fields $-|H_1|$ at the bottom and $+|H_1|$ at the top are applied. Using a generalized antiperiodic condition (GAPBC, see text) in y-direction, phase coexistence between two domains with opposite magnetization $\pm m_{\rm coex}$ in the bulk (symbolized by double arrows) is stabilized. The domains are separated by a domain wall inclined by an angle θ. **b** Schematic profiles of the layer magnetization $m^{(+)}(x)$, $m^{(-)}(x)$ across the film in the case of nonzero boundary fields. It is assumed (as in part **a**) that a state of partial wetting is realized. Fore more explanations, cf. text. From Block et al. [8.102]

for $|H_n| = -|H_1|$ which ensures that the inclined interface is strictly planar and not curved.

When one minimizes (8.61) with respect to θ (omitting the line tension term, which is negligible in the thermodynamic limit) one finds

$$\sigma(\theta)\cos\theta - \frac{\mathrm{d}\sigma(\theta)}{\mathrm{d}\theta}\sin\theta = \Delta f_1 \tag{8.62}$$

which is the well-know [8.146] modified Young's equation. A noted in (8.60), Δf_1 can be found straightforwardly from thermodynamic integration (in (8.60), one does not use the GAPBC needed to have an inclined interface, but rather has simple PBC in both y- and z-directions, and a state with uniform positive or negative ferromagnetic order). When $\sigma(\theta)$ is known, the contact angle can be extracted from (8.62). Of course, when $\sigma(\theta)$ is not known, the simulation using GAPBC with conserved total magnetization allows to estimate θ from direct "measurements" of the magnetization profile (as function of the variables y and x) in the system, as described in [8.102].

When F_{int} (8.61) and Δf_1 and $\gamma^{(+)}(|H_1|)$ hence have been found from thermodynamic integration methods, and θ is known as well as $\sigma(\theta)$, even the line tension τ can be estimated (τ is found to depend on both temperature T and contact angle θ, as expected). Of course, carrying out these procedures for different choices of the linear dimensions L_x, L_y and L_z one can check that the finite size effects (that occur for all interfacial properties, as discussed in Sect. 8.3) are small enough so that they can be neglected in this analysis. In the study of Block et al. [8.102] a parallel graphics processing units (GPU) implementation of the Ising model based on a checker board update was used, allowing to get very precise results for system sizes up to about 46 million Ising spins.

In the Ising model, where commonly the physical effect of the wall (apart from the missing neighbors) is described by the term $-H_1\sum_{i\in 1} S_i$ in the Hamiltonian, the relation $m_1 = -(\partial f_1/\partial H_1)_T$ for the surface layer magnetization $m_1 = \langle\sum_{i\in 1} S_i\rangle/N_1$ where N_1 is the number of spins in the 1st layer adjacent to the wall, has been a convenient starting point for the thermodynamic integration. However. it turns out that with a little thought the approach can be extended to off-lattice systems [8.44]. As a first step, consider an Ising system where the action of the surface magnetic field is not restricted to a single lattice plane but extends further into the bulk, described by a term $-H_1\sum_{i,k} S_{i,k} f(k)$, where the index $k = 1, 2, \ldots$ labels the planes parallel to the surface. Taking a derivative of the wall excess free energy f_1 then will generate a term $\tilde{m}_1 = -(\partial f_1/\partial H_1)_T$ with $\tilde{m}_1 = \frac{1}{N_1}\left\langle\sum_{i,k} S_{i,k} f(k)\right\rangle$ which also is straightforward to sample by standard Monte Carlo methods. Then $\Delta f_1 = f_1(H_1) - f_1(0) = -\int_0^{H_1}\tilde{m}_1(H_1')\mathrm{d}H_1'$. This approach then can be directly generalized to the off-lattice case, where the wall effect can be described by a potential $U_{\mathrm{wall}}(z) = U_{\mathrm{rep}}(z) - \varepsilon f_{\mathrm{att}}(z)$ with a repulsive part $U_{\mathrm{rep}}(z)$ that is infinity for $z = 0$ confining the particles in the region $z > 0$. The attractive part $U_{\mathrm{att}}(z)$ is described as a product of an energy parameter ε and a function $f_{\mathrm{att}}(z)$ describing its range. One then similarly can derive a relation [8.44]

$$\tilde{\rho}_\ell = -(\partial \sigma_{w\ell}/\partial \varepsilon)_T \quad , \quad \tilde{\rho}_\ell = \frac{1}{A} \int \mathrm{d}x \int \mathrm{d}y \int \mathrm{d}z F_{\mathrm{att}}(z) \left\langle \rho\left(x, y, z\right) \right\rangle \quad . \tag{8.63}$$

where A is the area of the wall. Equation (8.63) can then be used for a thermodynamic integration approach to extract differences of the wall tension $\sigma_{w\ell}$ of the liquid.

Also the ensemble switch method is a powerful approach to extract wall excess free energies [8.33, 8.34]. The basic equation is again (8.51): but now we mix one system ($\kappa = 0$) with periodic boundary conditions throughout and another one ($\kappa = 1$) with two equivalent walls at $z = 0$ and $z = L_z$. Both systems thus have exactly the same volume $V = L \times L \times L_z$ (in x, y directions periodic boundary conditions are used in both systems) and the same number of particles. Thus, (8.51) is a well-defined Hamiltonian for arbitrary values of κ also in this case, but the resulting surface tension $\sigma(L, L_z)$, (8.52), now is a wall tension. It has been found that in this case it is crucial to vary L_z over a wide range and carry out a linear extrapolation of the results for $\sigma(L, L_z)$ against $1/L_z$. This problem can be understood from the fact that in the systems with the walls in general there occurs an excess of particles near the walls (there is a positive surface excess density ρ_s for sufficiently strongly attractive potentials, and a negative surface excess density for repulsive ones). Due to this effect, the density in the center of the system will differ from the bulk density ρ_b that is used in the system without any walls by a correction of order ρ_s/L_z. But one can choose L relatively small, at least in cases where no critical wetting occurs, since the perturbation of the fluid near the wall does not involve large-scale fluctuations parallel to the wall, unlike interfaces between coexisting phases.

This method (and related earlier methods [8.147–8.149]) have been widely used to study the wall tension of hard spheres and other models of colloidal fluids [8.33, 8.34, 8.47, 8.140]; often the motivation has been to provide "benchmarks" against which the accuracy of analytical theories such as the density functional theory can be tested. Note that the standard method, based on the anisotropy of the pressure tensor analogous to (8.54) (see e.g. [8.40]) cannot be used when the interparticle potential is singular.

Putting then the knowledge of accurately determined interface tensions and wall tensions together, one can return to the problem of wetting phenomena in off-lattice fluids (Fig. 8.1). Pioneering studies along such lines have been performed both for the AO model of colloid-polymer mixtures [8.33, 8.34] and the Lennard-Jones fluid [8.150]. Grzelak and Errington [8.150] suggested to locate the prewetting transitions (Fig. 8.1b) for their model, demonstrating finite size effects of order L^{-2}, as expected from (8.50). A particularly compelling aspect of this study is the combination of results obtained from simulations in both the canonical and the grandcanonical ensemble. It was demonstrated that the finite size effects are of the same order (L^{-2}) but different magnitude in both ensembles, but do converge to the same result for $L \to \infty$. This study was based on the use of "transfer matrix Monte Carlo" [8.151], which is a variant of the extended ensemble methods described in Chap. 6, in order

to accurately locate the first-order prewetting transition in the parameter space of this model.

8.5 Discussion

While in the early days of the first applications of the Monte Carlo method to problems in the physics of solids and liquids the emphasis was on bulk properties, and a study of interfacial problems anyway would have been elusive due to the lack of computer power, in the last decades a shift of interest to the properties of inhomogeneous systems with surfaces and interfaces has occurred. E.g., when one is interested in the properties of nanoparticles, ultrathin films, etc., the power of analytical methods is clearly even more restricted than with respect to the properties of bulk condensed matter, and correspondingly the need for well-controlled Monte Carlo simulation approaches is evident.

In the present chapter,we hence have presented some extensions of the methodology that was exposed in previous chapters in order to give the reader a flavor of how one can obtain quantities such as interface tensions, wall excess free energies, contact angles and the singularities associated with wetting transitions from Monte Carlo simulations. As in previous chapters, we have focused on simple models such as the Ising/lattice gas model or the Lennard-Jones fluid, disregarding completely important questions such as "multiscale simulations" (where one tries to create bridges from the quantum mechanical description of electronic structure up to macroscopic properties of composite materials or of complex fluids such as biological matter) completely. The latter subject is clearly a main stream area of research (e.g. [8.152]), but the methodology described in the present chapter will be an indispensable input for this ambitious approach. We here have pointed out that the presence of interfaces and surfaces provides a substantial complication for properly performing a simulation and correctly analyzing the results: due to the naturally occurring anisotropy, directions parallel and perpendicular to the interface (and/or walls) are no longer equivalent, and in general it is no longer appropriate to choose the linear dimensions parallel (L) and perpendicular (L_z) to the interface (and/or the wall of a container) equivalent, and often size effects with respect to both L and L_z need to be carefully analyzed. For this purpose, both the statistical physics of interfaces and of wetting phenomena needs to be understood and taken into account, as attempted in the present chapter. Since Monte Carlo sampling does not give free energies directly, estimating the excess free energies due to interfaces and walls is a subtle matter, and many methods to achieve this task were derived. A few of them, which were found particularly useful, were described in this chapter.

Appendix

Throughout the guide we have encountered a number of algorithms. It was intended that everyone who actively participated in the guide would be able to try out the presented algorithms. All the algorithms given in the text were given in the pseudoprogramming language. We realize that at the moment not everybody is familiar with an algorithmic language. A guide, however, is of no use when the participants cannot sit down and do the simulations by themselves.

To help to ease the conversion from the algorithmic language into C we include in this appendix C versions of two of the basic algorithms. The participant will by now have a feeling for how some of the problems are attacked and will find no difficulty in adapting the given C programs to other algorithms given in the text.

A.1 Algorithm for the Random Walk Problem

The first listing is a program for the simple sampling of random walks.

```
/* ------------------------------------------------------------------------- */
/*                                                                           */
/*                              Algorithm                                    */
/*                                                                           */
/*        Simple Sampling for the Random Walk Problem                        */
/*                                                                           */
/* ------------------------------------------------------------------------- */

/* ==== Include files ==== */

#include <stdio.h>
#include <stdlib.h>
#include <math.h>


int main(int argc, char **argv) {

/*-------------------------------------------------------------------------*/
```


K. Binder and D. W. Heermann, *Monte Carlo Simulation in Statistical Physics*,
Graduate Texts in Physics, https://doi.org/10.1007/978-3-030-10758-1

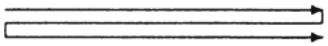
```
/*                          Main program                                      */
/*----------------------------------------------------------------------------*/

    /* ==== General declarations ==== */

    int   ip;
    int seed;

    long nwalk,n;

    long i,j;

    double x,y;
    double edx,edy,edxs,edys,xsqr,ysqr;

    /* ==== initialize ==== */

    seed  = 1234;
    n     = 1000;
    nwalk = 100000;

    edx   = 0.0;
    edy   = 0.0;

    edxs  = 0.0;
    edys  = 0.0;

    /*  ==== set up the random number generator ==== */

    srand(seed);

    /*------------------------------------------------------------ */
    /*               M o n t e  C a r l o                          */
    /*------------------------------------------------------------ */

    for(i=0; i<nwalk; i++) {

        x = 0.0;
        y = 0.0;

        for(j=0; j<n; j++) {

            ip = rand() % 4;
            switch (ip) {
                case 0: x -= 1.0;
                    break;
                case 1: y -= 1.0;
                    break;
                case 2: y += 1.0;
                    break;
                case 3: x += 1.0;
                    break;
            }

        }

        /* ==== accumulate the result ==== */

        xsqr = x * x;
        ysqr = y * y;

        edx  += xsqr;
        edy  += ysqr;

        edxs += xsqr * xsqr;
        edys += ysqr * ysqr;
    }

    /* ==== perform the averaging ==== */

    edx /= (float) nwalk;
    edy /= (float) nwalk;

    edxs /= (float) nwalk;
```

```
93      edys /= (float) nwalk;

95      printf("%f %f %f %f\n",edx,edy,edxs,edys);

97      exit(0);
} /* ---- end of main ---- */
```

A.2 Algorithm for Cluster Identification

The algorithm to identify all clusters in a given percolation configuration is based on an idea first proposed by Hoshen and Kopelman [A.1]. This idea was further developed by *Kertesz* (J. Kertesz, Private communication) and the algorithm described below makes extensive use of his development.

The algorithm uses labels which we also found useful for other algorithms given in the text. Suppose we scan the lattice from the upper left corner to the upper right corner and continue with the next row starting from the first site of the left edge:



etc.

This is rather like a typewriter working its way across the paper. We start by labeling the first occupied site with a "1". If the next site is occupied we carry the label "1" to this site. If we find another occupied site which is not connected with the cluster labeled "1", we label it with a "2", etc. Assume we are somewhere in the lattice. All the sites above and to left have been worked on. Assume further that the current site is occupied. We check whether the site above the current site is occupied. If so we read the label of this site. Next we check whether the site to the left is occupied. If so, we read the label of this site. What label should be assigned? The labels of the site above and the site to the left might not be the same. The two sites can belong to two, until now, unconnected clusters. The current site is a link between them. Both clusters have different labels. We need a means of identifying different labels with each other.

The idea used to equate different labels is to use a sort of permutation vector. If the cluster labeled, say, "10" becomes a part of the cluster labeled "7" we make an entry in element number "10" that we should really look for label "7".

That two clusters have to be recognized as one cluster can, of course, occur more than once. The cluster labeled "7" may be connected to cluster "5". How do we identify the root label?

Each time two clusters are identified with each other the number of sites in the cluster increases by more than one unit. From where do we get the information on the number of sites in a cluster?

The two questions raised in the preceding paragraphs can be answered together. We introduce two kinds of pointers. Positive integer numbers signify the number of

sites in the cluster. Negative integer numbers are the pointers to the next label. Going through the permutation vector we continue until a positive number is found.

A listing of a program where the idea outlined above is realized is given below.

```
/* ------------------------------------------------------------------------ */
/*                                                                          */
/*        Simple Sampling of the 2D percolation problem                     */
/*                                                                          */
/*        Analysis of the droplet distribution.                             */
/*                                                                          */
/* ------------------------------------------------------------------------ */

/* ==== Program constants ==== */

# define L          40

/* ==== Include files ==== */

#include <stdio.h>
#include <stdlib.h>
#include <math.h>


int main(int argc, char **argv)
{
    /* ------------------------------------------------------------------- */
    /*                         Declarations                                */
    /* ------------------------------------------------------------------- */


    int lattice[L][L];    /* 2D Ising spin system                  */
    int row[L+1];         /* temp                                  */
    int lptr[L*L+1];      /* Contains the pointer. The dimension   */
    /* is set, such that a magnetization of */
    /* one, i.e., one huge cluster is       */
    /* possible.                            */


    float clusterVector[L*L+1];
    float nsc[L*L+1];

    float averagePercProb;
    float avergageSuceptibility;
    float rlsqr;
    float chi;
    float r;
    float p;

    int nofs, irow, i, j;
    int la, cl, ms, up, mlarge, mcsmax;
    int seed, large;
    int mcs;
    int icol, left, mini, maxi;
    int nis, nup;
    int l_sqr;

    /* ------------------------------------------------------------------- */
    /*                         End of Declarations                         */
    /* ------------------------------------------------------------------- */

    /* ------------------------------------------------------------------- */
    /*                      Simulation parameters                          */
    /* ------------------------------------------------------------------- */


    mcsmax = 1000000;
    seed   = 12635;
    p      = 0.592746;


    /* ------------------------------------------------------------------- */
    /*                     Set and Initialize                              */
```

```
    /* ------------------------------------------------------------------- */

    /*  ==== set up the random number generator ==== */

    srand(seed);

    /* ==== Initialize the label array ==== */

    l_sqr  = L * L;
    rlsqr  = (float) l_sqr;
    large  = 0;                    /* largest cluster found. Also stack top */
    mlarge = 0;                    /* the largest cluster ever encountered  */

    averagePercProb       = 0.0;
    avergageSuceptibility = 0.0;

    for (i = 0; i <= l_sqr; i++){
        lptr[i]          = 0;
        nsc[i]           = 0.0;
        clusterVector[i] = 0.0;
    }

    maxi  = l_sqr + 2;

    /* ------------------------------------------------------------------- */
    /*                                                                     */
    /*             M O N T E   C A R L O   P A R T                         */
    /*                                                                     */
    /* ------------------------------------------------------------------- */

    for (mcs = 1; mcs <= mcsmax; mcs++){

        for (i = 0; i < L; i++){
            for (j = 0; j < L; j++){
                lattice[i][j] = 0;
                r = (float) rand() / (float) RAND_MAX;
                if (r < p) {
                    lattice[i][j] = 1;
                }
            }
        }

        /* ------------------------------------------------------------- */
        /*                                                               */
        /*             D r o p l e t   A n a l y s i s                   */
        /*                                                               */
        /* ------------------------------------------------------------- */

        /* ==== The array <<<row>>> holds always the previously ==== */
        /* ==== analysed row from the array lattice.            ==== */
        /* ==== First row is taken as entirely unoccupied, i.e. ==== */
        /* ==== free boundary conditions                        ==== */

        cl = 0;                /* Will be the largest cluster pointer */
        la = 0;

        for (irow = 0; irow <= L; irow++){
            row[irow] = maxi;
        }

        /* ==== Go through the lattice row by row ==== */

        for (irow = 0; irow < L; irow++){
            for (icol = 0; icol < L; icol++){
                if (lattice[irow][icol] != 1){
                    row[icol+1] = maxi;
                }
                else {
                    /*  ==== Check whether the spin is connected ==== */

                    up   = row[icol+1];
                    left = row[icol];

                    if (up != maxi){
```

```
                /* ==== Spin is connected to the previous row. ==== */

                if (lptr[up] < 0){
                    /* ==== Found a negative label, signaling a ==== */
                    /* ==== a pointer. Search now for the root. ==== */

                    ms = lptr[up];
                    while ( ms < 0 ){
                        la = -ms;
                        ms = lptr[la];
                    }
                    lptr[up] = -la;
                    up       = la;
                }
            }
```

```
            if (left != maxi){
                /* ==== Spin is connected to left neighbour. ==== */

                if (lptr[left] < 0) {
                    /* ==== Found a negative label, signaling a ==== */
                    /* ==== a pointer. Search now for the root. ==== */

                    ms = lptr[left];
                    while ( ms < 0 ){
                        la = -ms;
                        ms = lptr[la];
                    }
                    lptr[left] = -la;
                    left       = la;
                }
            }

            /*mini = min(up,left); */
            mini = (up < left) ? up : left;
            if (mini == maxi) {
                /* ==== Spin is not connected. Assign new label ==== */

                cl++;
                row[icol+1] = cl;
                lptr[cl]    = 1;

            }
            else {
                /* ==== Spin is connected. Find minimum label ==== */

                nofs = 1;
                if (up != left){
                    /* ==== Possibly two clusters joined      ==== */
                    /* ==== into one. Count both parts        ==== */
                    /* ==== towards the number of cluster.    ==== */

                    if (up != maxi){
                        nofs    += lptr[up];
                        lptr[up] = -mini;
                    }

                    if (left != maxi){
                        nofs      += lptr[left];
                        lptr[left] = -mini;
                    }
                }
                else {
                    /* ==== The two possible cluster are ==== */
                    /* ====  part o one. Count only one  ==== */

                    nofs      += lptr[left];
                    lptr[left] = -mini;
                }
                row[icol+1] = mini;
                lptr[mini]  = nofs;
            }
        }
```

```
216             }
            }
218
            /* ---------------------------------------------------------------- */
220         /*                                                                  */
            /*          E n d    D r o p l e t   A n a l y s i s                */
222         /*                                                                  */
            /* ---------------------------------------------------------------- */
224
226         /* ---------------------------------------------------------------- */
            /*                                                                  */
228         /*                 Analysis of droplet numbers                      */
            /*                                                                  */
230         /* ---------------------------------------------------------------- */

232         for (i = 0; i <= l_sqr; i++) {
                clusterVector[i] = 0.0;
234         }

236         /* ==== Get the largest droplet and the size distribution ==== */

238         large = 0;
            for (i = 1; i <= cl; i++){
240             nis = lptr[i];
                if ( nis > 0 ) {
242                 if ( nis > large ) { large = nis; }
                    clusterVector[nis] += 1.0;
244                 nsc[nis]           += 1.0;
                }
246         }

248         /* ==== Compute the susceptibility and accumulate ==== */

250         chi = 0.0;
            nup = 0;
252         for (i = 1; i <= large ; i++) {
                chi += i * i * clusterVector[i];
254             nup += i * clusterVector[i];
            }
256         chi -= large * large * clusterVector[large];

258         averagePercProb       += nup / rlsqr;
            avergageSuceptibility += chi / rlsqr;
260
            if (large > mlarge) { mlarge = large; }
262
            for (i = 0; i <= l_sqr; i++) {
264             lptr[i] = 0;
            }
266
            /* ---------------------------------------------------------------- */
268         /*                                                                  */
            /*               End analysis of droplet numbers                    */
270         /*                                                                  */
            /* ---------------------------------------------------------------- */
272

274     }  /* end of Monte Carlo loop */

276     /* ---------------------------------------------------------------- */
        /*                                                                  */
278     /*                 P r i n t    r e s u l t s                       */
        /*                                                                  */
280     /* ---------------------------------------------------------------- */

282     averagePercProb       /= (float) mcsmax;
        avergageSuceptibility /= (float) mcsmax;
284
        printf("percProb = %f\nchi       = %f\n",averagePercProb,
286                                  avergageSuceptibility);
        printf("\n***** Droplet Analysis *****\n");
288     printf("Largest droplet = %d\n",mlarge);
```

```
290     for (i = 1; i <= mlarge; i++){
            if (nsc[i] != 0.0) {
292             printf("%d  %f\n ",i,nsc[i]/ (float) mcsmax);
            }
294     }
        printf("\n");
296
} /* end of main program */
```

References

General References

K. Binder (ed.): *Applications of the Monte Carlo Method in Statistical Physics,* Topics Curr. Phys., Vol. 36, 2nd ed. (Springer, Berlin, Heidelberg, 1987)
D.W. Heermann: *Computer Simulation Methods in Theoretical Physics* (Springer, Berlin, Heidelberg, 1986)
See also [2.8, 2.10, 6.54, 6.61, 6.95]

Chapter 2

[2.1] K. Kremer, A. Baumgärtner, K. Binder, J. Phys. A **15**, 2879 (1982)
[2.2] P.J. Flory, *Principles of Polymer Chemistry* (Cornell University Press, Ithaca, NY, 1953)
[2.3] P.G. de Gennes, *Scaling Concepts in Polymer Physics* (Cornell University Press, Ithaca, NY, 1979)
[2.4] A. Baumgärtner, Topics current physics, in *Applications of the Monte Carlo Method in Statistical Physics*, vol. 36, 2nd edn., ed. by K. Binder (Springer, Berlin, Heidelberg, 1987) (Chap. 5)
[2.5] A. Milchev, K. Binder, D.W. Heermann, Z. Phys. B **63**, 521 (1986)
[2.6] L.D. Landau, E.M. Lifshitz, *Statistical Physics* (Pergamon, Oxford, 1958)
[2.7] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, E. Teller, J. Chem. Phys. **21**, 1087 (1953)
[2.8] K. Binder (ed.), Topics current physics, in *Monte Carlo Methods in Statistical Physics*, vol. 7, 2nd edn. (Springer, Berlin, Heidelberg, 1986)
[2.9] H. Müller-Krumbhaar, K. Binder, J. Stat. Phys. **8**, 1 (1973)
[2.9a] S. Duane, A.D. Kennedy, B.J. Pendleton, D. Roweth, Phys. Lett. B **195**, 216 (1987); D.W. Heermann, P. Nielaba, M. Rovere, Comput. Phys. Commun. **60**, 311 (1990); B. Mehlig, D.W. Heermann, B.M. Forrest, Phys. Rev. B **45**, 679 (1992)
[2.10] K. Binder, in *Phase Transitions and Critical Phenomena*, vol. 5b, ed. by C. Domb, M.S. Green (Academic, New York, 1976), p. 1
[2.11] R.J. Glauber, J. Math. Phys. **4**, 294 (1963)
[2.12] K. Kawasaki, in *Phase Transitions and Critical Phenomena*, vol. 2, ed. by C. Domb, M.S. Green (Academic, New York, 1972), p. 443

K. Binder and D. W. Heermann, *Monte Carlo Simulation in Statistical Physics*,
Graduate Texts in Physics, https://doi.org/10.1007/978-3-030-10758-1

[2.13] A. Milchev, D.W. Heermann, K. Binder, J. Stat. Phys. **44**, 749 (1986); see also A.D. Bruce, J. Phys. A **18**, L873 (1985)
[2.14] A.K. Kron, Vysokomol. Soedin. **7**, 1228 (1965) [Polym. Sci. USSR **7**, 1361 (1965)]; F.T. Wall, F. Mandel, J. Chem. Phys. **63**, 4592 (1975)
[2.15] P.E. Rouse, J. Chem. Phys. **21**, 1272 (1953)
[2.16] P.H. Verdier, W.H. Stockmayer, J. Chem. Phys. **36**, 227 (1962); P.H. Verdier, J. Chem. Phys. **45**, 2122 (1966); P.H. Verdier, J. Chem. Phys. **59**, 6119 (1973); D.E. Kranbuehl, P.H. Verdier, J. Chem. Phys. **71**, 2662 (1979); H.J. Hilhorst, J.M. Deutch, J. Chem. Phys. **63**, 5153 (1975); H. Boots, J.M. Deutch, J. Chem. Phys. **67**, 4608 (1977)
[2.17] A. Baumgärtner, K. Binder, J. Chem. Phys. **71**, 2541 (1979); J. Chem. Phys. **75**, 2994 (1981); A. Baumgärtner, Annu. Rev. Phys. Chem. **35**, 419 (1984)
[2.18] W. Oed, Appl. Inform. **7**, 358 (1982)
[2.19] A.B. Bortz, M.H. Kalos, J.L. Lebowitz, J. Comput.Phys. **17**, 10 (1975); see also K. Binder [2.8, p. 1]
[2.20] A. Sadiq, J. Comput. Phys. **55**, 387 (1984)
[2.21] H. Müller-Krumbhaar, K. Binder, Z. Phys. **254**,269 (1972); K. Binder, H. Müller-Krumbhaar, Phys. Rev. B **7**, 3297 (1973)
[2.22] For a review see W. Selke, in *Alloy Phase Stability*, ed. by A. Gonis, L.M. Stocks (Elsevier, New York, 1988)
[2.23] M.E. Fisher, in *Critical Phenomena*, ed. by M.S. Green (Academic, New York, 1971); M.E. Fisher, M.N. Barber, Phys. Rev. Lett. **28**, 1516 (1972)
[2.24] M.N. Barber, in *Phase Transitions and Critical Phenomena*, vol. 8, ed. by C. Domb, J.L. Lebowitz (Academic, New York, 1983) (Chap. 2)
[2.25] K. Binder, Z. Phys. B **43**, 119 (1981); K. Binder, Phys. Rev. Lett. **47**, 693 (1981)
[2.26] K. Binder, Ferroelectrics **73**, 43 (1987)
[2.27] D.P. Landau, Phys. Rev. B **13**, 2997 (1976); D.P. Landau, Phys. Rev. B **14**, 255 (1976)
[2.28] S. Wansleben, Comput. Phys. Commun. **43**, 315 (1987)
[2.29] K. Binder, D.P. Landau, Phys. Rev. Lett. **52**, 318 (1984); Surf. Sci. **151**, 409 (1985)
[2.30] K. Binder, Phys. Status Solidi **46**, 567 (1971); Physica **62**, 508 (1972); Z. Angew. Phys. **30**, 51 (1970)
[2.31] V. Wildpaner, Z. Phys. **270**, 215 (1974)
[2.32] For a review of "simulations of small systems" see H. Müller-Krumbhaar, in [2.8, Chap. 5]
[2.33] K. Binder, H. Rauch, V. Wildpaner, J. Phys. Chem. Solids **31**, 391 (1970)
[2.34] K. Binder, D.P. Landau, J. Appl. Phys. **57**, 3306 (1985); K. Binder, D.P. Landau, D.M. Kroll, J. Magn. Magn. Mater. **54–57**, 669 (1986); K. Binder, D.P. Landau, D.M. Kroll, Phys. Rev. Lett. **56**, 2272 (1986)
[2.35] K. Binder, D.P. Landau, Phys. Rev. B **37**, 1745 (1988)
[2.36] I.M. Kim, D.P. Landau, Surf. Sci. **110**, 415 (1981)
[2.37] K. Binder, Thin Solid Films **20**, 367 (1974)
[2.38] K. Binder, P.C. Hohenberg, Phys. Rev. B **9**,2194 (1974); K.K. Mon, K. Binder, D.P. Landau, Phys. Rev. B **35**, 3683 (1987)
[2.39] H. Müller-Krumbhaar, K. Binder, Int. J. Magn. **3**, 113 (1972)
[2.40] K. Binder, M.H. Kalos, in [2.8, Chap. 6]; K.W. Kehr, K. Binder, Topics current physics, in *Applications of the Monte Carlo Method in Statistical Physics*, ed. by K. Binder, vol. 36, 2nd edn. (Springer, Berlin, Heidelberg, 1987) (Chap. 6)
[2.41] K. Binder, D. Stauffer, in [2.8, Chap. 8]; K. Binder, A.P. Young, Rev. Mod. Phys. **58**, 867 (1986)
[2.42] H.J. Herrmann, Phys. Rep. **136**, 143 (1986)
[2.43] A. Baumgärtner, K. Binder, K. Kremer, Faraday Symp. Chem. Soc. **18**, 37 (1983); A. Baumgärtner, J. Polym. Sci. C Symp. **73**, 181 (1985)
[2.44] H.-O. Carmesin, K. Binder, Europhys. Lett. **4**,269 (1987); Z. Phys. B **68**, 375 (1987)
[2.45] J.D. Gunton, M. San Miguel, P.S. Sahni, in *PhaseTransitions and Critical Phenomena*, vol. 8, ed. by C. Domb, J.L. Lebowitz (Academic, New York, 1983) (Chap. 3); K. Binder, in *Condensed Matter Research Using Neutrons*, ed. by S.W. Lovesey, R. Scherm (Plenum,

New York, 1984), p. 1; K. Binder, in *Alloy Phase Stability*, ed. by A. Gonis, L.M. Stocks (Nijhoff, Dordrecht, 1990); K. Binder, D.W. Heermann, in *Scaling Phenomena in Disordered Systems*, ed. by R. Pynn, A. Skjeltorp (Plenum, New York, 1987), p. 207

[2.46] A. Sadiq, K. Binder, J. Stat. Phys. **35**, 617 (1984); K. Binder, Ber. Bunsenges. Phys. Chem. **90**, 257 (1986); G.S. Grest, P.S. Sahni, Phys. Rev. B **30**, 2261 (1984); E.T. Gawlinski, M. Grant, J.D. Gunton, K. Kaski, Phys. Rev. B **31**, 281 (1985); K. Binder, D.W. Heermann, A. Milchev, A. Sadiq, in*Heidelberg Colloquium on Glassy Dynamics*, ed. by L. van Hemmen, I. Morgenstern, Lecture Notes in Physics, vol. 275 (Springer, Berlin, Heidelberg, 1987), p. 154

[2.47] A. Sadiq, K. Binder, Surf. Sci. **128**, 350 (1985)

[2.48] For background information on the master equation description of stochastic Ising models and related models see [2.12]

[2.49] N. Madras, A.D. Sokal, J. Stat. Phys. **50**, 109 (1988)

[2.50] A. Sariban, K. Binder, Macromolecules **21**, 711 (1988)

[2.51] This time is discussed by Binder and Young [2.41b], by Binder et al. [2.46e] and by R.G. Palmer, Adv. Phys. **31**, 669 (1982)

[2.52] M.Y. Choi, B.A. Huberman, Phys. Rev. B **29**, 2796 (1984)

[2.53] Compare, e.g., Algorithm 3 described in [2.8, p. 33]

[2.54] R. Kutner, K. Binder, K.W. Kehr, Phys. Rev.B **26**, 2967 (1982); Phys. Rev. **28**, 1846 (1983)

[2.55] K.W. Kehr, R. Kutner, K. Binder, Phys. Rev. B **23**, 4931 (1981)

[2.56] Recent reviews of the molecular dynamics method arefound in W.G. Hoover, *Molecular Dynamics*. Lecture Notes in Physics, vol. 258 (Springer, Berlin, Heidelberg, 1986); and D.W. Heermann, *Computer Simulation Methods in Theoretical Physics*(Springer, Berlin, Heidelberg, 1986) (Chap. 2)

[2.57] R. Friedberg, J.E. Cameron, J. Chem. Phys. **52**, 6049 (1970)

[2.58] K.E. Schmidt, Phys. Rev. Lett. **51**, 2175 (1983); G. Parisi, in *Progress in Gauge Field Theory*, ed. by G. t'Hooft et al. (Plenum, New York, 1984), p. 531; G.G. Batrouni, G. Katz, A. Kronfeld, G. Lepage, S. Svetitsky, K. Wilson, Phys.Rev. D **32**, 2736 (1985); J. Goodman, A. Sokal, Phys. Rev. Lett. **56**, 1015 (1986); R.H. Swendsen, J.S. Wang, Phys. Rev. Lett. **57**, 2607 (1986); S. Duane, Nucl. Phys. B **257** [FS14], 652 (1985); J. Kogut, Nucl. Phys. B **275** [FS17], 1 (1986); E. Dagotto, J. Kogut, Phys. Rev. Lett. **58**, 299 (1987)

[2.59] P.C. Hohenberg, B.I. Halperin, Rev. Mod. Phys. **49**, 435 (1977)

[2.60] D. Stauffer, Phys. Rep. **54**, 1 (1979); *Introduction to Percolation Theory* (Taylor and Francis, London, 1985)

[2.61] Cluster counting algorithms have been described by H. Müller-Krumbhaar [2.8, Chap. 5]; J. Hoshen, R. Kopelman, Phys. Rev. B **14**, 3438 (1976); by D. Stauffer, in *Disordered Systems and Localization*. Lecture Notes in Physics, vol. 149 (Springer, Berlin, Heidelberg, 1981), p. 9; and in [2.62]. A different approachis the cluster growth method [P.L. Leath, Phys. Rev. B **14**, 5046 (1976)]

[2.62] K. Binder, D. Stauffer, in *Applications of the Monte Carlo Method in Statistical Physics*, 2nd edn., ed. by K. Binder (Springer, Berlin, Heidelberg, 1987) (Chap. 8)

[2.63] D.W. Heermann, D. Stauffer, Z. Phys. B **40**, 133 (1980)

[2.64] S. Kirkpatrick, in *Ill-Condensed Matter*, ed. by R. Balian, R. Maynard, G. Toulouse (North-Holland, Amsterdam, 1979), p. 321

[2.65] A.P. Young, S. Kirkpatrick, Phys. Rev. B **15**, 440 (1982)

[2.66] E. Stoll, K. Binder, T. Schneider, Phys. Rev. B **8**, 3266 (1973)

[2.67] K. Binder, H. Rauch, Z. Phys. **219**, 201 (1969)

[2.68] D. Stauffer, K. Binder, Z. Phys. B **30**, 313(1978); **41**, 237 (1981)

[2.69] H.E. Stanley, *An Introduction to Phase Transitions and Critical Phenomena* (Oxford University Press, Oxford, 1971)

[2.70] K. Binder, Phys. Rev. A **25**, 1699 (1982)

[2.71] K.K. Kaski, K. Binder, J.D. Gunton, Phys. Rev. B **29**, 3996 (1984)

[2.72] M.N. Barber, R.B. Pearson, J.L. Richardson, D. Touissaint, Phys. Rev. B **32**, 1720 (1985)

[2.73] D.P. Landau, J. Magn. Magn. Mater. **31–34**, 11 (1983)

[2.74] C.S.S. Murty, D.P. Landau, J. Appl. Phys. **55**, 2429 (1984); M.S.S. Challa, D.P. Landau, Phys. Rev. B **33**, 437 (1986)
[2.75] D.P. Landau, K. Binder, Phys. Rev. B **31**, 5946 (1985)
[2.76] K. Binder, M. Nauenberg, V. Privman, A.P. Young, Phys.Rev. B **31**, 1498 (1985); K. Binder, Z. Phys. B **61**, 13 (1985)
[2.77] R.N. Bhatt, A.P. Young, Phys. Rev. Lett. **54**, 924 (1985)
[2.78] B. Dünweg, K. Binder, Phys. Rev. B **36**, 6935 (1987)
[2.79] G.S. Pawley, R.H. Swendsen, D.J. Wallace, K.G. Wilson, Phys. Rev. B **29**, 4030 (1984)
[2.80] G. Bhanot, D. Duke, R. Salvador, Phys. Rev.B **33**, 7841 (1986); J. Stat. Phys. **44**, 985 (1986)
[2.81] G. Parisi, F. Rapuano, Phys. Lett. B **157**, 301 (1985)
[2.82] A. Hoogland, A. Compagner, H.W.J. Blöte, Phys. A **132**, 457 (1985)
[2.83] K. Binder, D.P. Landau, Phys. Rev. B **21**, 1941 (1980)
[2.84] D.P. Landau, Phys. Rev. B **27**, 5604 (1983)
[2.85] S. Krinsky, D. Mukamel, Phys. Rev. B **16**, 2313 (1977)
[2.86] R.H. Swendsen, S. Krinsky, Phys. Rev. Lett. **43**, 177 (1979)
[2.87] J. Oitmaa, J. Phys. A **14**, 1159 (1981)
[2.88] M.P. Nightingale, Phys. Lett. A **59**, 486 (1977)
[2.89] M. Nauenberg, B. Nienhuis, Phys. Rev. Lett. **33**, 944 (1974)
[2.89a] A.M. Ferrenberg, D.P. Landau, Phys. Rev. B **44**, 5081 (1991)
[2.90] M.N. Barber, W. Selke, J. Phys. A **15**, L617 (1982)
[2.91] S. Miyashita, H. Takano, Prog. Theor. Phys. **73**, 1122 (1985)
[2.92] S. Wansleben, D.P. Landau, J. Appl. Phys. **61**, 3968 (1987); D.P. Landau, S.Y. Tang, S. Wansleben, J. Phys. (Paris) **49**, Colloque **8**, 1525 (1989)
[2.93] Z. Racz, Phys. Rev. B **13**, 263 (1976); M.E. Fisher, Z. Racz, Phys. Rev. B **13**, 5039 (1976)
[2.94] R. Kretschmer, K. Binder, D. Stauffer, J. Stat. Phys. **15**, 267 (1976)
[2.95] K. Binder, Z. Phys. B **45**, 61 (1981)
[2.96] For a review on the roughening transition see J.D. Weeks, in *Ordering in Strongly Fluctuating Condensed Matter Systems*, ed. by T. Riste (Plenum, New York, 1980), p. 293
[2.97] For a general review on first-order phase transitions see K. Binder, Rep. Prog. Phys. **50**, 783 (1987)
[2.98] Y. Imry, Phys. Rev. B **21**, 2042 (1980)
[2.99] M.E. Fisher, A.N. Berker, Phys. Rev. B **26**, 2507 (1982)
[2.100] H.W.J. Blöte, M.P. Nightingale, Phys. A **112**, 405 (1982); J.L. Cardy, M.P. Nightingale, Phys. Rev. B **27**, 4256 (1983); P. Kleban, C.-K. Hu, Bull. Am. Phys. Soc. **27**, 92(A) (1982)
[2.101] V. Privman, M.E. Fisher, J. Stat. Phys. **33**, 385 (1983)
[2.102] K. Binder, D.P. Landau, Phys. Rev. B **30**, 1477 (1984)
[2.103] M.E. Fisher, V. Privman, Phys. Rev. B **32**, 447 (1985)
[2.104] M.S.S. Challa, D.P. Landau, K. Binder, Phys. Rev. B **34**, 1841 (1986) An error of [2.104] concerning the weights $a_{\pm}$ was corrected by C. Borgs, R. Kotecky, J. Stat. Phys. **61**, 79 (1990)
[2.105] C.N. Yang, Phys. Rev. **85**, 808 (1951)
[2.106] R.B. Potts, Proc. Camb. Philos. Soc. **48**, 106 (1952); for a review see F.Y. Wu, Rev. Mod. Phys. **54**, 235 (1982)
[2.107] T. Kihara, Y. Midzuno, T. Shizume, J. Phys. Soc. Jpn. **9**, 681 (1954); R.J. Baxter, J. Phys. C **6**, L445 (1973)

Chapter 3

[3.1] D. Knuth, *The Art of Computer Programming*, vol. 2 (Addison-Wesley, Reading, MA, 1969)

[3.2] D.W. Heermann, *Computer Simulation Methods in Theoretical Physics* (Springer, Berlin, Heidelberg, 1986)
[3.3] M.P. Allen, D.J. Tildesley, *Computer Simulation of Liquids* (Clarendon, Oxford, 1987)
[3.4] M.H. Kalos, P.A. Whitlock, *Monte Cario Methods*, vol. 1 (Wiley, New York, 1986)
[3.5] J.H. Ahrens, U. Dieter, *Pseudo Random Numbers* (Wiley, New York) (in preparation)
[3.6] D.H. Lehmer, in *Proceedings of the 2nd Symposium on Large-Scale Digital Computing Machinery*, vol. 142 (Harvard University Press, Cambridge, MA, 1951)
[3.7] A. Milchev, K. Binder, D.W. Heermann, Z. Phys. B **63**, 521 (1986)
[3.8] H.E. Stanley, J. Stat. Phys. **36**, 843 (1984)
[3.9] L. Onsager, Phys. Rev. **65**, 117 (1944)

Chapter 4

[4.1] R.H. Swendsen, J.-S. Wang, Phys. Rev. Lett. **58**, 86 (1987)
[4.2] U. Wolff, Phys. Rev. Lett. **60**, 1461 (1988)
[4.3] U. Wolff, Nucl. Phys. B **300**, 501 (1988)
[4.4] F. Niedermayer, Phys. Rev. Lett. **61**, 2026 (1988)
[4.5] R.G. Edwards, A.D. Sokal, Phys. Rev. D **38**, 2009 (1988)
[4.6] R.G. Edwards, A.D. Sokal, Phys. Rev. D **40**, 1374 (1989)
[4.7] U. Wolff, Phys. Rev. Lett. **62**, 361 (1989)
[4.8] U. Wolff, Phys. Lett. B **228**, 379 (1989)
[4.9] U. Wolff, Nucl. Phys. B **322**, 759 (1989)
[4.10] W. Klein, T.S. Ray, P. Tamayo, Phys. Rev. Lett. **62**, 163 (1989)
[4.11] T.S. Ray, P. Tamayo, W. Klein, Phys. Rev. A **39**, 5949 (1989)
[4.12] C. Frick, K. Jansen, P. Seuferling, Phys. Rev. Lett. **63**, 2613 (1989)
[4.13] X.-J. Li, A.D. Sokal, Phys. Rev. Lett. **63**, 827 (1989)
[4.14] C.-K. Hu, K.-S. Mak, Phys. Rev. B **40**, 5007 (1989)
[4.15] J.-S. Wang, D. Chowdhury, J. Phys. (Paris) **50**, 2905 (1989)
[4.16] J.-S. Wang, Phys. A **161**, 249 (1989)
[4.17] Z. Alexandrowicz, Phys. A **160**, 310 (1989)
[4.18] D. Stauffer, Phys. A **162**, 27 (1989)
[4.19] A.N. Burkitt, D.W. Heermann, Europhys. Lett. **10**, 207 (1989)
[4.20] A. Coniglio, F. Liberto, G. Monroy, F. Peruggi, J. Phys. A Math. Gen. **22**, L837 (1989)
[4.21] R.C. Brower, P. Tamayo, Phys. Rev. Lett. **62**, 1087 (1989)
[4.22] J.-S. Wang, R.H. Swendsen, R. Kotecky, Phys. Rev. Lett. **63**, 109 (1989)
[4.23] J.-S. Wang, R.H. Swendsen, R. Kotecky, Phys. Rev. B **42**, 2465 (1990)
[4.24] W. Janke, Phys. Lett. A **148**, 306 (1990)
[4.25] M. D'Onorio de Meo, D.W. Heermann, K. Binder, J. Stat. Phys. **60**, 585 (1990)
[4.26] J.-S. Wang, D. Stauffer, Z. Phys. B **78**, 145 (1990)
[4.27] T.S. Ray, P. Tamayo, J. Stat. Phys. **60**, 851 (1990)
[4.28] T.S. Ray, J.-S. Wang, Phys. A **167**, 580 (1990)
[4.29] J.-S. Wang, R.H. Swendsen, Phys. A **167**, 565 (1990)
[4.30] D. Kandel, R. Ben-Av, E. Domany, Phys. Rev. Lett. **65**, 941 (1990)
[4.31] R. Ben-Av, D. Kandel, E. Katznelson, P.G. Lauwers, J. Stat. Phys. **58**, 125 (1990)
[4.32] R.C. Brower, S. Huang, Phys. Rev. D **62**, 708 (1990)
[4.33] J.-S. Wang, M. Wöhlert, H. Mühlenbein, D. Chowdhury, Phys. A **166**, 173 (1990)
[4.34] J.-S. Wang, W. Selke, V.S. Dotsenko, V.B. Andreichenko, Europhys. Lett. **11**, 301 (1930); Phys. A **164**, 221 (1990)
[4.35] F. Niedermayer, Phys. Lett. B **237**, 473 (1990)
[4.36] D.W. Heermann, A.N. Burkitt, Phys. A **162**, 210 (1990)
[4.37] J.-S. Wang, Phys. A **164**, 240 (1990)

[4.38] P. Tamayo, R.C. Brower, W. Klein, J. Stat. Phys. **58**, 1083 (1990)
[4.39] M. Hasenbusch, Nucl. Phys. B **333**, 581 (1990)
[4.40] M. Hasenbusch, S. Meyer, Phys. Lett. B **241**, 238 (1990)
[4.41] M. Hasenbusch, S. Meyer, Phys. Rev. Lett. **66**, 530 (1991)
[4.42] X.-J. Li, A.D. Sokal, Phys. Rev. Lett. **67**, 1482 (1991)
[4.43] L.-J. Chen, C.-K. Hu, K.-S. Mak, Comput. Phys. Commun. **66**, 377 (1991)
[4.44] D. Kandel, E. Domany, Phys. Rev. B **43**, 8539 (1991)
[4.45] R.C. Brower, P. Tamayo, Phys. A **193**, 314 (1992)
[4.46] R.H. Swendsen, J.-S. Wang, A.M. Ferrenberg, Topics in applied physics, in *The Monte Carlo Method in Condensed Matter Physics*, vol. 71, 2nd edn., ed. by K. Binder (Springer, Berlin, Heidelberg, 1995) (Chap. 4)
[4.47] U. Wolff, in *Computational Methods in Field Theory*, ed. by H. Gausterer, C.B. Lang (Springer, Berlin, Heidelberg, 1992), p. 125
[4.48] A.D. Sokal, in *Quantum Fields on the Computer*, ed. by M. Creutz (World Scientific, Singapore, 1992), p. 211
[4.49] H.G. Evertz, M. Marcu, G. Lana, Phys. Rev. Lett. **70**, 875 (1993); S. Wiseman, E. Domany, Phys. Rev. E **48**, 4080 (1983)
[4.50] P.D. Caddington, L. Han, Phys. Rev. B **50**, 3058 (1994)
[4.51] E. Luijten, H.W.J. Blöte, Int. J. Mod. Phys. C **6**, 359 (1995); Phys. Rev. Lett. **76**, 1557, 3662 (1996)
[4.52] J. Machta, Y.S. Choi, A. Lucke, T. Schweizer, L.V. Chayes, Phys. Rev. Lett. **75**, 2792 (1995)
[4.53] T.B. Liverpool, S.C. Glotzer, Phys. Rev. E **53**, R 4255 (1996)
[4.54] J.E. Gubernatis, N. Kawashima, in *Monte Carloand Molecular Dynamics of Condensed Matter Systems,* ed. by K. Binder, G. Ciccotti (Società Italiana di Fisica, Bologna, 1996), p. 518; N. Kawashima, J.E. Gubernatis, Phys. Rev. E **51**, 1547 (1995)
[4.55] P.W. Kasteleyn, P.M. Fortuin, J. Phys. Soc. Jpn. **26**(Suppl.), 11 (1969); C.M. Fortuin, P.W. Kasteleyn, Physica **57**, 536 (1972)
[4.56] J. Goodman, A.D. Sokal, Phys. Rev. Lett. **56**, 1015 (1986); Phys. Rev. D **40**, 2035 (1989)
[4.57] D. Kandel, E. Domany, A. Brandt, Phys. Rev. B **40**, 330 (1989)
[4.58] M. Hasenbusch, S. Meyer, G. Mack, Nucl. Phys. B (Proc. Suppl.) **20**, 110 (1991)
[4.59] W. Janke, T. Sauer, Chem. Phys. Lett. **201**, 499 (1993)
[4.60] W. Janke, Springer proceedings in physics, in *Computer Simulation Studies in Condensed Matter Physics VII*, vol. 78, ed. by D.P. Landau, K.K. Mon, H.-B. Schüttler (Springer, Berlin, Heidelberg, 1994), p. 29
[4.61] W. Janke, T. Sauer, J. Stat. Phys. **78**, 75 (1995)
[4.62] G.A. Baker Jr., N. Kawashima, Phys. Rev. Lett. **75**, 994 (1995)
[4.63] A.M. Ferrenberg, R.H. Swendsen, Phys. Rev. Lett. **61**, 2635 (1988)
[4.64] Z.W. Salsburg, D. Jacobson, W. Fickett, W.W. Wood, J. Chem. Phys. **30**, 65 (1959)
[4.65] D.A. Chestnut, Z.W. Salsburg, J. Chem. Phys. **38**, 2861 (1963); I.R. McDonald, K. Singer, Discuss. Faraday Soc. **43**, 40 (1967)
[4.66] A.M. Ferrenberg, R.H. Swendsen, Phys. Rev. Lett. **63**, 1195 (1989)
[4.67] B.A. Berg, Int. J. Mod. Phys. C **3**, 1083 (1992)
[4.68] B. Dünweg, in *Monte Carlo and Molecular Dynamics of Condensed Matter Systems*, ed. by K. Binder, G. Ciccotti (Società Italiana di Fisica, Bologna, 1996), p. 215
[4.69] D. Frenkel, B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications* (Academic, San Diego, 1996)
[4.70] C.H. Bennett, J. Comput. Phys. **22**, 245 (1976)
[4.71] J.P. Valleau, D.N. Card, J. Chem. Phys. **57**, 5457 (1972)
[4.72] G.M. Torrie, J.P. Valleau, J. Comput. Phys. **23**, 187 (1977); Chem. Phys. Lett. **28**, 578 (1974)
[4.73] J.P. Valleau, G.M. Torrie, in *Statistical Mechanics, Part A: Equilibrium Techniques*, ed. by B.J. Berne (Plenum, New York, 1977)
[4.74] B.A. Berg, T. Neuhaus, Phys. Rev. Lett. **68**, 9 (1992)

[4.75] U. Hansmann, B.A. Berg, T. Neuhaus, Int. J. Mod. Phys. C **3**, 1155 (1992)
[4.76] B.A. Berg, U. Hansmann, T. Neuhaus, Phys. Rev. B **47**, 497 (1993)
[4.77] B.A. Berg, U. Hansmann, T. Neuhaus, Z. Phys. B **90**, 229 (1993)
[4.78] E. Marinari, G. Parisi, Europhys. Lett. **19**, 457 (1992)
[4.79] A.P. Lyubartsev, A.A. Martsinovski, S.V. Sherkenov, P.N. Vorontsov-Velyaminov, J. Chem. Phys. **96**, 1776 (1992)
[4.80] P.M.C. de Oliveira, T.S.P. Penna, H.J. Herrmann, Braz. J. Phys. **26**, 677 (1996); P.M.C. de Oliveira, T.J.P. Penna, H. Heermann, Eur. Phys. J. B **1**, 205 (1998); P.M.C. de Oliveira, Eur. Phys. J. B **6**, 111 (1998); J.-S. Wang, Eur. Phys. J. B **8**, 287 (1998); J.-S. Wang, L.-W. Lee, Comput. Phys. Commun. **127**, 131 (2000)
[4.81] J. Lee, Phys. Rev. Lett. **71**, 211 (1993)
[4.82] N.A. Alves, B.A. Berg, R. Villanova, Phys. Rev. B **41**, 383 (1990)
[4.83] G. Bhanot, S. Black, P. Carter, R. Salvador, Phys.Lett. B **183**, 331 (1987); G. Bhanot, R. Salvador, S. Black, R. Toral, Phys. Rev. Lett. **59**, 803 (1987)
[4.84] W. Janke, S. Kappler, Phys. Rev. Lett. **74**, 212 (1995)
[4.85] N.B. Wilding, A.D. Bruce, J. Phys. Condens. Matter **4**, 3087 (1992)
[4.86] N.B. Wilding, Z. Phys. B **93**, 113 (1993)
[4.87] N.B. Wilding, M. Müller, J. Chem. Phys. **102**, 2562 (1995)
[4.88] N.B. Wilding, Phys. Rev. E **52**, 602 (1995)
[4.89] M. Müller, N.B. Wilding, Phys. Rev. E **51**, 2079 (1995)
[4.90] N.B. Wilding, M. Müller, K. Binder, J. Chem. Phys. **105**, 802 (1996)
[4.91] N.B. Wilding, K. Binder, Phys. A **231**, 439 (1996)
[4.92] N.B. Wilding, P. Nielaba, Phys. Rev. E **53**, 926 (1996)
[4.93] N.B. Wilding, *Annual Reviews of Computational Physics IV*, ed. by D. Stauffer (World Scientific, Singapore, 1996), p. 37
[4.94] N.B. Wilding, Springer proceedings in physics, in *Computer Simulation Studies in Condensed Mattter VIII*, vol. 80, ed. by D.P. Landau, K.K. Mon, H.B. Schüttler (Springer, Berlin, Heidelberg, 1995), p. 18
[4.95] E. Riedel, F.J. Wegner, Z. Phys. **225**, 195 (1969)
[4.96] M.E. Fisher, Rev. Mod. Phys. **46**, 597 (1974)
[4.97] K. Binder, H.P. Deutsch, Europhys. Lett. **18**, 667 (1992)
[4.98] H.P. Deutsch, K. Binder, J. Phys. II (Fr.) **3**, 1049 (1993)
[4.99] K.K. Mon, K. Binder, Phys. Rev. E **48**, 2498 (1993)
[4.100] E. Luijten, H.W.J. Blöte, K. Binder, Phys. Rev.E **54**, 4626 (1996); E. Luijten, H.W.J. Blöte, K. Binder, Phys. Rev. E **56**, 6540 (1997); E. Luijten, K. Binder, Phys. Rev. E **58**, R4060 (1998)
[4.101] K. Binder, R. Evans, D.P. Landau, A.M. Ferrenberg, Phys. Rev. E **53**, 5023 (1996)
[4.102] A. Coniglio, W. Klein, J. Phys. A **13**, 2775 (1980)
[4.103] C.K. Hu, Phys. Rev. B **29**, 5103 (1984)
[4.104] M.E. Fisher, Physics **3**, 267 (1967)
[4.105] K. Binder, Ann. Phys. **98**, 390 (1976)
[4.106] K. Binder, J.-S. Wang, J. Stat. Phys. **55**, 87 (1989)
[4.107] J.L. Cardy, *Phase Transitions and Critical Phenomena*, vol. 11 (Academic, New York, 1987), p. 55
[4.108] D.P. Landau, S. Tang, S. Wansleben, J. Phys. (Paris) Colloq. C8, **49**, 1525 (1988)
[4.109] A.M. Ferrenberg, D.P. Landau, Y.J. Wong, Phys. Rev. Lett. **69**, 3382 (1992)
[4.110] K. Binder, in *Monte Carlo and Molecular Dynamics Simulations in Polymer Science*, ed. by K. Binder (Oxford University Press, New York, 1995), p. 356
[4.111] H.-P. Deutsch, J. Stat. Phys. **67**, 1039 (1992)
[4.112] H.-P. Deutsch, K. Binder, Macromolecules **25**, 6214 (1992)
[4.113] A. Sariban, K. Binder, J. Chem. Phys. **86**, 5859 (1987)
[4.114] I. Carmesin, K. Kremer, Macromolecules **21**, 2819 (1988); H.-P. Deutsch, K. Binder, J. Chem. Phys. **94**, 2294 (1991)

[4.115] K. Binder, *Festkörperprobleme (Advances in Solid State Physics)*, vol. 26 (Vieweg, Braunschweig, 1986), p. 133
[4.116] H.-P. Deutsch, K. Binder, Macromol. Chem. Macromol. Symp. **65**, 59 (1993)
[4.117] H.-P. Deutsch, J. Chem. Phys. **99**, 4825 (1993)
[4.118] M. Müller, K. Binder, Macromolecules **28**, 1825 (1995)
[4.119] J.J. Rehr, N.D. Mermin, Phys. Rev. A **8**, 472 (1973)
[4.120] A.Z. Panagiotopolous, Mol. Phys. **61**, 813 (1987); Mol. Simul. **9**, 1 (1992)
[4.121] A.Z. Panagiotopolous, in *Supercritical Fluids-Fundamentals for Application*, ed. by E. Kiran, J.M.H. Levelt-Sengers (Kluwer, Dordrecht, 1994)
[4.122] D. Frenkel, B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications* (Academic, San Diego, 1996)
[4.123] M.P. Allen, in *Monte Carlo and Molecular Dynamics of Condensed Matter Systems*, ed. by K. Binder, G. Ciccotti (Società Italiana di Fisica, Bologna, 1996), p. 255
[4.124] K. Binder, D.P. Landau, Phys. Rev. B **13**, 1140 (1976)
[4.125] R.B. Stinchcombe, *Phase Transitions and Critical Phenomena*, vol. 7 (Academic, London, 1983), p. 151
[4.126] T. Nattermann, J. Villain, Phase Transit. **11**, 5 (1988)
[4.127] W. Selke, L.N. Shchur, A.L. Talapov, *Annual Reviews of Computational Physics*, vol. 1 (World Scientific, Singapore, 1994), p. 17
[4.128] H. Rieger, Phys. Rev. B **52**, 6659 (1995)
[4.129] M. D'Onorio de Meo, J.D. Reger, K. Binder, Phys. A **220**, 628 (1995)
[4.130] V. Pereyra, P. Nielaba, K. Binder, Z. Phys. B **97**, 197 (1995)
[4.131] K. Eichhorn, K. Binder, Europhys. Lett. **30**, 331 (1995); J. Phys. Condens. Matter **8**, 5209 (1996)
[4.132] I.D. Lawrie, S. Sarbach, *Phase Transitions and Critical Phenomena*, vol. 9 (Academic, London, 1984), p. 1
[4.133] V.L. Ginzburg, Sov. Phys. Solid State **2**, 1824 (1960)
[4.134] P.G. de Gennes, J. Phys. (Paris) Lett. **38**, L-441 (1977); J.F. Joanny, J. Phys. A **11**, L-117 (1978); K. Binder, J. Chem. Phys. **79**, 6387 (1983)
[4.135] G. Kamieniarz, H.W.J. Blöte, J. Phys. A **26**, 201 (1993)
[4.136] E. Brézin, J. Zinn-Justin, Nucl. Phys. B **257**, 867 (1985)
[4.137] M.B. Bouabci, C.E.I. Carneiro, Phys. Rev. **1354**, 359 (1996)
[4.138] M. Blume, V.J. Emery, R.B. Griffiths, Phys. Rev. A **4**, 4946 (1976)

Chapter 5

[5.1] C. Rickwardt, P. Nielaba, K. Binder, M. Müser, Phys. Rev. B **63**, 045204 (2001)
[5.2] P. Nielaba, *Annual Reviews of Computational Physics V*, ed. by D. Stauffer (World Scientific, Singapore, 1997), p. 137
[5.3] R. Martonak, W. Paul, K. Binder, Phys. Rev. E **57**, 2425 (1998)
[5.4] D. Marx, H. Wiechert, Adv. Chem. Phys. **95**, 213 (1996)
[5.5] D. Marx, O. Opitz, P. Nielaba, K. Binder, Phys. Rev. Lett. **79**, 2908 (1993)
[5.6] R.P. Feynman, A.R. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York, 1965)
[5.7] H.F. Trotter, Proc. Am. Math. Soc. **10**, 545 (1959)
[5.8] M. Suzuki, Progr. Theor. Phys. **46**, 1337 (1971)
[5.9] B.J. Berne, D. Thirumalai, Ann. Rev. Phys. Chem. **37**, 401 (1986)
[5.10] J.D. Doll, J.E. Gubernatis (eds.), *Quantum Simulations* (World Scientific, Singapore, 1990)
[5.11] D.M. Ceperley, Rev. Mod. Phys. **67**, 279 (1995)
[5.12] D.M. Ceperley, in *Monte Carlo and Molecular Dynamics of Condensed Matter Systems*, ed. by K. Binder, G. Ciccotti (Società Italiana di Fisica, Bologna, 1996), p. 445

[5.13] M.H. Müser, P. Nielaba, K. Binder, Phys. Rev. B **51**, 2723 (1995)
[5.14] D.N. Batchelder, D.L. Losee, R.O. Simmons, Phys. Rev. **73**, 873 (1966)
[5.15] R.P. Feynman, Phys. Rev. **90**, 1116 (1953); ibid **91**, 1291, 1301 (1953)
[5.16] M. Suzuki (ed.), *Quantum Monte Carlo Methods* (Springer, Berlin, Heidelberg, New York, 1986)
[5.17] M. Suzuki (ed.), *Quantum Monte Carlo Methods in Condensed Matter Physics* (World Scientific, Singapore, 1992)
[5.18] H. de Raedt, W. von der Linden, in *The Monte Carlo Method in Condensed Matter Physics*, ed. by K. Binder (Springer, Berlin, Heidelberg, New York, 1992), p. 249
[5.19] M. Suzuki, Commun. Math. Phys. **51**, 183 (1976)
[5.20] H. de Raedt, A. Lagendijk, Phys. Rep. **127**, 233 (1985)
[5.21] M. Suzuki, Progr. Theor. Phys. **56**, 1454 (1976)
[5.22] M. Barma, B.S. Shastry, Phys. Rev. B **18**, 3351 (1978)
[5.23] D.P. Landau, K. Binder, *A Guide to Monte Carlo Simulations in Statistical Physics* (Cambridge University Press, Cambridge, 2000) (Chap. 8)
[5.24] J.E. Hirsch, R.L. Sugar, D.J. Scalapino, R. Blancenbecler, Phys. Rev. B **26**, 5033 (1982)
[5.25] J. Hubbard, Proc. R. Soc. A **276**, 238 (1963)
[5.26] J.E. Hirsch, Phys. Rev. B **28**, 4059 (1983)
[5.27] J.E. Hirsch, Phys. Rev. B **31**, 4403 (1985)
[5.28] J.E. Gubernatis, N. Kawashima, in *Monte Carlo and Molecular Dynamics of Condensed Matter Systems*, ed. by K. Binder, G. Ciccotti (Società Italiana di Fisica, Bologna, 1996), p. 519
[5.29] D. Marx, M.H. Müser, J. Phys. Condens. Matter **11**, R117 (1999)

Chapter 6

[6.1] L.D. Landau, E.M. Lifshitz, *Statistical Physics*, 3rd edn. (Pergamon Press, Oxford, 1980)
[6.2] K. Binder, Z. Phys. B **45**, 61 (1981)
[6.3] S. Kämmerer, B. Dünweg, K. Binder, M. D'Onorio De Meo, Phys. Rev. B **53**, 2345 (1996)
[6.4] W. Schweika, D.P. Landau, K. Binder, Phys. Rev. B **53**, 8937 (1996)
[6.5] F.F. Haas, F. Schmid, K. Binder, Phys. Rev. B **61**, 15077 (2000)
[6.6] R. Lipowsky, J. Appl. Phys. **55**, 2485 (1984)
[6.7] M. Schrader, *Diplomarbeit* (Johannes-Gutenberg-Universität Mainz, 2009, unpublished)
[6.8] M. Schrader, P. Virnau, K. Binder, Phys. Rev. E **79**, 061104 (2009)
[6.9] P. Virnau, M. Müller, J. Chem. Phys. **120**, 10925 (2004)
[6.10] L.G. Mac Dowell, P. Virnau, M. Müller, K. Binder, J. Chem. Phys. **120**, 5293 (2004)
[6.11] L.G. Mac Dowell, V.K. Shen, J.R. Errington, J. Chem. Phys. **125**, 034705 (2006)
[6.12] M. Müller, J.J. de Pablo, in *Computer Simulations in Condensed Matter: From Materials to Chemical Biology*, vol. 1, ed. by M. Ferrario, G. Ciccotti, K. Binder (Springer, Berlin, Heidelberg, New York, 2006), p. 67
[6.13] H. Furukawa, K. Binder, Phys. Rev. A **26**, 556 (1982)
[6.14] B.A. Berg, U. Hansmann, T. Neuhaus, Z. Phys. B **90**, 229 (1993)
[6.15] J.E. Hunter, W.P. Reinhardt, J. Chem. Phys. **103**, 8627 (1995)
[6.16] K. Binder, M.H. Kalos, J. Stat. Phys. **22**, 363 (1980)
[6.17] M. Biskup, L. Chayes, R. Kotecky, Europhys. Lett. **60**, 21 (2002)
[6.18] K. Binder, Phys. A **319**, 99 (2003)
[6.19] E.A. Carlen, M.C. Carvalho, R. Esposito, J.L. Lebowitz, R. Marra, Mol. Phys. **103**, 3141 (2005)
[6.20] P. Virnau, M. Müller, L.G. Mac Dowell, K. Binder, New J. Phys. **6**, 7 (2004). http://stacks.iop.org/1367-2630/6/7
[6.21] K. Binder, Phys. Rev. A **25**, 1699 (1982)

[6.22] K. Binder, M. Müller, W. Oed, J. Chem. Soc. Faraday Trans. **91**, 2369 (1995)
[6.23] M. Müller, M. Schick, J. Chem. Phys. **105**, 8885 (1996)
[6.24] M. Müller, L.G. Mac Dowell, Macromolecules **33**, 3902 (2000)
[6.25] P. Virnau, M. Müller, L.G. Mac Dowell, K. Binder, J. Chem. Phys. **121**, 2169 (2004)
[6.26] R.L.C. Vink, J. Horbach, K. Binder, Phys. Rev. E **71**, 011401 (2005)
[6.27] B.M. Mognetti, L. Yelash, P. Virnau, W. Paul, K. Binder, M. Müller, L.G. Mac Dowell, J. Chem. Phys. **128**, 104501 (2008)
[6.28] C. Dellago, P.G. Bolhuis, P.L. Geissler, Adv. Chem. Phys. **123**, 1 (2002)
[6.29] W. Janke (ed.), *Rugged Free Energy Landscapes: Common Computational Approaches to Spin Glasses, Structural Glasses and Biological Macromolecules*. Lecture Notes in Physics, vol. 736 (Springer, Berlin, Heidelberg, New York, 2008)
[6.30] K. Binder, A.P. Young, Rev. Mod. Phys. **58**, 801 (1986)
[6.31] K. Binder, W. Kob, *Glassy Materials and Disordered Solids. An Introduction to Their Statistical Mechanics* (World Scientific, Singapore, 2005)
[6.32] U.H.E. Hansmann, Y. Okamoto, in *Annual Reviews in Computational Physics VI*, ed. by D. Stauffer (World Scientific, Singapore, 1999), p. 129
[6.33] Y. Sugita, Y. Okamoto, in *Lecture Notes in Computational Science and Engineering*, ed. by T. Schlick, H.H. Gan (Springer, Berlin, Heidelberg, New York, 2002), p. 304
[6.34] P.G. Bolhuis, in *Computer Simulations in Condensed Matter: From Materials to Chemical Biology*, ed. by M. Ferrario, G. Ciccotti, K. Binder (Springer, Berlin, Heidelberg, New York, 2006), p. 393
[6.35] M. Bachmann, W. Janke, in [6.29], p. 203
[6.36] A. Irbäck, in [6.29], p. 269
[6.37] U.H.E. Hansmann, in [6.29], p. 293
[6.38] Y. Sugita, A. Mitsutake, Y. Okamoto, in [6.29], p. 369
[6.39] J. Luettmer-Strathmann, F. Rampf, W. Paul, K. Binder, J. Chem. Phys. **128**, 064903 (2008)
[6.40] K. Binder, W. Paul, T. Strauch, F. Rampf, V. Ivanov, J. Luettmer-Strathmann, J. Phys. Condens. Matter **20**, 494215 (2008)
[6.41] A.P. Young (ed.), *Spin Glasses and Random Fields* (World Scientific, Singapore, 1998)
[6.42] A.P. Young, in *Computer Simulations in Condensed Matter Systems: From Materials to Chemical Biology, Vol 2*, ed. by M. Ferrario, G. Ciccotti, K. Binder. Lecture Notes in Physics, vol. 704 (Springer, Berlin, Heidelberg, New York, 2006), p. 31
[6.43] A. Billoire, in [6.29], p. 11
[6.44] K. Binder, C. Brangian, W. Kob, in [6.29], p. 47
[6.45] T. Aspelmeier, A. Billoire, E. Marinari, M.A. Moore, J. Phys. A Math. Theor. **41**, 324008 (2008)
[6.46] A.P. Young, J. Phys. A Math. Theor. **41**, 324016 (2008)
[6.47] G.M. Torrie, J.P. Valleau, Chem. Phys. Lett. **28**, 578 (1974)
[6.48] G.M. Torrie, J.P. Valleau, J. Comput. Phys. **23**, 187 (1977)
[6.49] B.A. Berg, T. Neuhaus, Phys. Lett. B **267**, 249 (1991)
[6.50] B.A. Berg, T. Neuhaus, Phys. Rev. Lett **68**, 9 (1992)
[6.51] B.A. Berg, Int. J. Mod. Phys. C **4**, 249 (1993)
[6.52] W. Janke, Phys. A **254**, 164 (1998)
[6.53] B.A. Berg, Comput. Phys. Commun. **153**, 397 (2003)
[6.54] B.A. Berg, *Markov Chain Monte Carlo Simulations and Their Statistical Analysis* (World Scientific, Singapore, 2004)
[6.55] M. Bachmann, W. Janke, J. Chem. Phys. **120**, 6779 (2004)
[6.56] F. Wang, D.P. Landau, Phys. Rev. Lett. **86**, 2050 (2001)
[6.57] F. Wang, D.P. Landau, Phys. Rev. E **64**, 056101 (2001)
[6.58] D.P. Landau, F. Wang, Comput. Phys. Commun. **147**, 674 (2002)
[6.59] B.J. Schulz, K. Binder, M. Müller, D.P. Landau, Phys. Rev. E **67**, 067102 (2003)
[6.60] S. Trebst, D.A. Huse, M. Troyer, Phys. Rev. E **70**, 046701 (2004)
[6.61] D.P. Landau, K. Binder, *A Guide to Monte Carlo Simulation in Statistical Physics*, 3rd edn. (Cambridge University Press, Cambridge, 2009)

[6.62] C. Zhou, R.N. Bhatt, Phys. Rev. E **72**, 025701 (2005)
[6.63] D.P. Landau, in [6.29], p. 353
[6.64] N.B. Wilding, in *Computer Simulations in Condensed Matter: From Materials to Chemical Biology, Vol 1*, ed. by M. Ferrario, G. Ciccotti, K. Binder. Lecture Notes in Physics, vol. 703 (Springer, Berlin, Heidelberg, New York, 2006), p. 39
[6.65] A.P. Lyubartsev, A.A. Matsinovski, S.V. Shevkunov, P.N. Vorontsov-Velyaminov, J. Chem. Phys. **96**, 1776 (1992)
[6.66] E. Marinari, G. Parisi, Europhys. Lett. **19**, 451 (1992)
[6.67] J. Lee, Phys. Rev. Lett. **71**, 211 (1993); Erratum ibid. **71**, 2353 (1993)
[6.68] K. Hukushima, H. Takayama, K. Nemoto, Int. Mod. Phys. C **3**, 337 (1996)
[6.69] P.M.C. de Oliveira, T.J.P. Penna, H.J. Herrmann, Braz. J. Phys. **26**, 677 (1996)
[6.70] P.M.C. de Oliveira, T.J.P. Penna, H.J. Herrmann, Eur. Phys. J. B **1**, 205 (1998)
[6.71] J.S. Wang, L.W. Lee, Comput. Phys. Commun. **127**, 131 (2000)
[6.72] Y. Iba, Int. J. Mod. Phys. C **12**, 623 (2001)
[6.73] Q. Yan, F. Faller, J.J. de Pablo, J. Chem. Phys. **116**, 8745 (2002)
[6.74] J.S. Wang, R.H. Swendsen, J. Stat. Phys. **106**, 245 (2002)
[6.75] I. Carmesin, K. Kremer, Macromolecules **21**, 2819 (1988)
[6.76] H.P. Deutsch, K. Binder, J. Chem. Phys. **94**, 2294 (1991)
[6.77] C. Dellago, P.G. Bolhuis, F.S. Csajka, D. Chandler, J. Chem. Phys. **108**, 1964 (1998)
[6.78] C. Dellago, P.G. Bolhuis, D. Chandler, J. Chem. Phys. **108**, 9236 (1998)
[6.79] P.G. Bolhuis, C. Dellago, D. Chandler, Faraday Discuss. **110**, 421 (1998)
[6.80] P.G. Bolhuis, D. Chandler, C. Dellago, P.L. Geissler, Ann. Rev. Phys. Chem. **53**, 291 (2002)
[6.81] C. Dellago, D. Chandler, in *Bridging Time Scales: Molecular Simulation for the Next Decade*, ed. by P. Nielaba, M. Mareschal, G. Ciccotti (Springer, Berlin, Heidelberg, New York, 2002)
[6.82] C. Dellago, in *Handbook of Materials Modeling*, ed. by S. Yip (Springer, Berlin, Heidelberg, New York, 2005), p. 1585
[6.83] C. Dellago, P.G. Bolhuis, P.L. Geissler, in *Computer Simulations in Condensed Matter. From Materials to Chemical Biology*, vol 1., ed. by M. Ferrario, G. Ciccotti, K. Binder (Springer, Berlin, Heidelberg, New York, 2006), p. 349
[6.84] W. Ren, E. Vanden-Eijnden, Phys. Rev. B **66**, 05301 (2002)
[6.85] A. Laio. M. Parrinello, Proc. Natl. Acad. Sci. U.S.A. **99**, 12562 (2002)
[6.86] R. Elber, A. Ghosh, A. Gardenas, H. Stern, Adv. Chem. Phys. **126**, 93 (2004)
[6.87] C. Micheletti, A. Laio, M. Parrinello, Phys. Rev. Lett. **92**, 170601 (2004)
[6.88] E. Weinan, W. Ren, E. Vanden-Eijnden, J. Phys. Chem. B **109**, 6688 (2005)
[6.89] G. Bussi, A. Laio, M. Parrinello, Phys. Rev. Lett. **96**, 090601 (2006)
[6.90] A. Laio, M. Parrinello, in *Computer Simulations in Condensed Matter: From Materials to Chemical Biology*, vol. 1, ed. by M. Ferrario, G. Cicotti, K. Binder (Springer, Berlin, Heidelberg, New York, 2006), p. 315
[6.91] H. Müller-Krumbhaar, K. Binder, J. Stat. Phys. **8**, 1 (1973)
[6.92] K. Binder, H. Müller-Krumbhaar, Phys. Rev. B **9**, 2328 (1974)
[6.93] K. Binder, D. Stauffer, Adv. Phys. **25**, 343 (1976)
[6.94] A.C. Pan, D. Chandler, J. Phys. Chem. B **108**, 19681 (2004)
[6.95] D. Frenkel, B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications*, 2nd edn. (Academic Press, San Diego, 2002)
[6.96] C.H. Bennett, J. Comput. Phys. **22**, 245 (1976)
[6.97] B.M. Mognetti, P. Virnau, L. Yelash, W. Paul, K. Binder, M. Müller, L.G. MacDowell, J. Chem. Phys. (2009, in press)
[6.98] F.Y. Wu, Rev. Mod. Phys. **54**, 235 (1982)
[6.99] K. Binder, D.P. Landau, Phys. Rev. B **30**, 1477 (1984)
[6.100] C. Borgs, R. Kotecky, J. Stat. Phys. **61**, 79 (1990)
[6.101] A.M. Ferrenberg, R.H. Swendsen, Phys. Rev. Lett. **61**, 2635 (1988)
[6.102] A.M. Ferrenberg, R.H. Swendsen, Phys. Rev. Lett. **63**, 1195 (1989)
[6.103] B.A. Berg, in [6.29], p. 317

[6.104] B.A. Berg, J. Stat. Phys. **82**, 323 (1996)
[6.105] B.A. Berg, A. Billoire, W. Janke, Phys. Rev. E **61**, 12143 (2000)
[6.106] B.J. Schulz, K. Binder, M. Müller, Phys. Rev. E **71**, 046705 (2005)
[6.107] A.M. Ferrenberg, D.P. Landau, K. Binder, Phys. Rev. E **58**, 3353 (1998)
[6.108] K. Vollmayr, J.D. Reger, M. Scheucher, K. Binder, Z. Phys. B Condens. Matter **91**, 113 (1993)
[6.109] T. Neuhaus, S. Hager, J. Stat. Phys. **113**, 47 (2003)
[6.110] D. Moroni, P.R. ten Wolde, P.G. Bolhuis, Phys. Rev. Lett. **94**, 235703 (2005)
[6.111] A.C. Zettelmoyer, *Nucleation* (Marcel Dekker, New York, 1969)
[6.112] Z.W. Salsburg, J.D. Jacobson, W.S. Fickett, W.W. Wood, J. Chem. Phys. **30**, 65 (1959)
[6.113] S.-H. Tsai, F. Wang, D.P. Landau, H.-Q. Lin, Phys. Rev. E **75**, 061108 (2007)
[6.114] Z. Chen, Y. Gu, J. Bioinform. Comput. Biol. **4**, 317 (2006)
[6.115] E. Luijten, H.W.J. Blöte, Int. J. Mod. Phys. C **6**, 359 (1995)
[6.116] J. Lui, E. Luijten, Phys. Rev. E **71**, 066701 (2005)
[6.117] Y.C. Kim, M.E. Fisher, Phys. Rev. E **68**, 041506 (2003)
[6.118] J.C. Grossman, J. Chem. Phys. **117**, 1434 (2002)
[6.119] M. Nekovee, W.M.C. Foulkes, R.J. Needs, Math. Comput. Simul. **62**, 463 (2003)
[6.120] W. Schattke, R. Bahnsen, R. Redmer, Progr. Surf. Sci. **72**, 87 (2003)
[6.121] S.-I. Lui, J. Chem. Phys. **121**, 10365 (2004)

Chapter 7

[7.1] D.G. Kendall, Random fluctuations in the age-distribution of a population whose development is controlled by the simple "birth-and-death" process **12**(2), 278–285 (1950)
[7.2] A.B. Bortz, M.H. Kalos, J.L. Lebowitz, A new algorithm for monte carlo simulation of ising spin systems. J. Comput. Phys. **17**(1), 10–18 (1975)
[7.3] D.T. Gillespie, A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. J. Comput. Phys. **22**(4), 403–434 (1976)
[7.4] M.A. Gibson, J. Bruck, Efficient exact stochastic simulation of chemical systems with many species and many channels. J. Phys. Chem. A **104**(9), 1876–1889, 03 (2000)
[7.5] T. Fricke, J. Schnakenberg, Monte-carlo simulation of an inhomogeneous reaction-diffusion system in the biophysics of receptor cells. Zeitschrift für Physik B Condensed Matter **83**(2), 277–284 (1991)
[7.6] D.T. Gillespie, Approximate accelerated stochastic simulation of chemically reacting systems. J. Chem. Phys. **115**(4), 1716–1733 (2001)
[7.7] Y. Cao, D.T. Gillespie, L.R. Petzold, Avoiding negative populations in explicit Poisson tau-leaping. J. Chem. Phys. **123**(5), 054104 (2005)
[7.8] T.P. Schulze, Efficient kinetic Monte Carlo simulation. J. Comput. Phys. 0021-9991 **227**(4), 2455–2462 (2008)
[7.9] W.M. Young, E.W. Elcock, Monte Carlo studies of vacancy migration in binary ordered alloys: I. Proc. Phys. Soc. **89**(3), 735 (1966)
[7.10] D.T. Gillespie, Exact stochastic simulation of coupled chemical reactions. J. Phys. Chem. **81**(25), 2340–2361, 12 (1977)
[7.11] D.R. Cox, H.D. Miller, *The Theory of Stochastic Processes* (Taylor and Francis, 1977)
[7.12] A.F. Voter, *Introduction to the Kinetic Monte Carlo Method* (Springer Netherlands, Dordrecht, 2007), pp. 1–23
[7.13] A. Chatterjee, D.G. Vlachos, An overview of spatial microscopic and accelerated kinetic monte carlo methods. J. Comput. Aided Mater. Des. **14**(2), 253–308 (2007)
[7.14] D.T Gillespie, Stochastic simulation of chemical kinetics. Annu. Rev. Phys. Chem. **58**, 35–55 (2007)

[7.15] H. Li, Y. Cao, L.R. Petzold, D.T. Gillespie, Algorithms and software for stochastic simulation of biochemical reacting systems. Biotechnol. Prog. **24**(1), 56–61 (2008)
[7.16] A.P.J. Jansen, *An Introduction to Kinetic Monte Carlo Simulations of Surface Reactions*, vol. 856 (Springer, Berlin, Heidelberg)
[7.17] Y. Cao, H. Li, L. Petzold, Efficient formulation of the stochastic simulation algorithm for chemically reacting systems. J. Chem. Phys. **121**(9), 4059–4067 (2004)
[7.18] J.M. McCollum, G.D. Peterson, C.D. Cox, M.L. Simpson, N.F. Samatova, The sorting direct method for stochastic simulation of biochemical systems with varying reaction execution behavior. Comput. Biol. Chem. **30**(1), 39–49, 2 (2006)
[7.19] B. Meng, W.H. Weinberg, Dynamical monte carlo studies of molecular beam epitaxial growth models: interfacial scaling and morphology. Surf. Sci. **364**(2), 151–163 (1996)
[7.20] V.V. Bulatov, W. Cai, *Computer Simulations of Dislocations*. Oxford Series on Materials Modelling (Oxford University Press, 2006)
[7.21] S.A. Baeurle, T. Usami, A.A. Gusev, A new multiscale modeling approach for the prediction of mechanical properties of polymer-based nanomaterials. Polymer **47**(26), 8604–8617, 12 (2006)
[7.22] K. Binder, M.H. Kalos, Springer topics in current physics, in *Monte Carlo Methods in Statistical Physics*, vol. 7, p. 225, ed. by K. Binder (Springer, Berlin, Heidelberg, 1979)
[7.23] J.L. Blue, I. Beichl, F. Sullivan, Faster Monte Carlo simulations. Phys. Rev. E **51**(2), R867–R868, 02 (1995)
[7.24] T.P. Schulze, Kinetic Monte Carlo simulations with minimal searching. Phys. Rev. E **65**(3), 036704, 02 (2002)
[7.25] A. Violi, A.F. Sarofim, G.A. Voth, Kinetic Monte Carlo–molecular dynamics approach to model soot inception. Combust. Sci. Technol. **176**(5–6), 991–1005, 05 (2004)
[7.26] M.A. Katsoulakis, A. Chatterjee, D.G. Vlachos, Int. J. Multiscale Comput. Eng. **3**(135) (2005)
[7.27] Aleksandar Donev, Vasily V. Bulatov, Tomas Oppelstrup, George H. Gilmer, Babak Sadigh, Malvin H. Kalos, A first-passage kinetic Monte Carlo algorithm for complex diffusion-reaction systems. J. Comput. Phys. **229**(9), 3214–3236 (2010)
[7.28] K.A. Fichthorn, W.H. Weinberg, Theoretical foundations of dynamical Monte Carlo simulations. J. Chem. Phys. **95**(2), 1090–1096 (1991)
[7.29] S.A. Serebrinsky, Physical time scale in kinetic Monte Carlo simulations of continuous-time Markov chains. Phys. Rev. E **83**(3), 037701, 03 (2011)
[7.30] P.A. Maksym, Fast Monte Carlo simulation of MBE growth. Semicond. Sci. Technol. **3**(6), 594 (1988)
[7.31] B.D. Lubachevsky, Efficient parallel simulations of dynamic Ising spin systems. J. Comput. Phys. **75**(1), 103–122 (1988)
[7.32] K.M. Chandy, J. Misra, Distributed simulation: a case study in design and verification of distributed programs. IEEE Trans. Softw. Eng. **SE-5**(5), 440–452 (1979)
[7.33] J. Misra, Distributed discrete-event simulation. ACM Comput. Surv. **18**, 39 (1986)
[7.34] Y. Shim, J.G. Amar, Semirigorous synchronous sublattice algorithm for parallel kinetic Monte Carlo simulations of thin film growth. Phys. Rev. B **71**(12), 125432, 03 (2005)
[7.35] G. Arampatzis, M.A. Katsoulakis, P. Plecháč, M. Taufer, X. Lifan, Hierarchical fractional-step approximations and parallel kinetic Monte Carlo algorithms. J. Comput. Phys. **231**(23), 7795–7814 (2012)
[7.36] I. Martin-Bragado, J. Abujas, P.L. Galindo, J. Pizarro, Synchronous parallel kinetic Monte Carlo: implementation and results for object and lattice approaches. Nucl. Instrum. Methods Phys. Res. Sect. B Beam Interact. Mater. At. **352**, 27–30 (2015)
[7.37] G. Korniss, Z. Toroczkai, M.A. Novotny, P.A. Rikvold, From massively parallel algorithms and fluctuating time horizons to nonequilibrium surface growth. Phys. Rev. Lett. **84**(6), 1351–1354, 02 (2000)
[7.38] G. Korniss, M.A. Novotny, H. Guclu, Z. Toroczkai, P.A. Rikvold, Suppressing roughness of virtual times in parallel discrete-event simulations. Science 299(5607), 677, 01 (2003)

[7.39] M. Kardar, G. Parisi, Y.-C. Zhang, Dynamic scaling of growing interfaces. Phys. Rev. Lett. **56**(9), 889–892, 03 (1986)
[7.40] L. Xu, M. Taufer, S. Collins, D.G. Vlachos, Parallelization of tau-leap coarse-grained Monte Carlo simulations on GPUs, in *2010 IEEE International Symposium on Parallel & Distributed Processing (IPDPS)* (2010), pp. 1–9
[7.41] J. Nielsen, M. d'Avezac, J. Hetherington, M. Stamatakis, Parallel kinetic Monte Carlo simulation framework incorporating accurate models of adsorbate lateral interactions. J. Chem. Phys. **139**(22), 224706 (2013)
[7.42] I. Komarov, R.M. D'Souza, Accelerating the gillespie exact stochastic simulation algorithm using hybrid parallel execution on graphics processing units. PLOS One **7**(11), e46693, 11 (2012)
[7.43] G. Korniss, M.A. Novotny, P.A. Rikvold, Parallelization of a dynamic Monte Carlo algorithm: a partially rejection-free conservative approach. J. Comput. Phys. **153**(2), 488–508 (1999)
[7.44] E. Martínez, P.R. Monasterio, J. Marian, Billion-atom synchronous parallel kinetic Monte Carlo simulations of critical 3D Ising systems. J. Comput. Phys. **230**(4), 1359–1369 (2011)
[7.45] W. Chen, E. De Schutter, Parallel steps: large scale stochastic spatial reaction-diffusion simulation with high performance computers. Front. Neuroinformatics **11**, 13 (2017)
[7.46] F. Jiménez, C.J. Ortiz, A GPU-based parallel object kinetic Monte Carlo algorithm for the evolution of defects in irradiated materials. Comput. Mater. Sci. **113**, 178–186 (2016)
[7.47] G. Klingbeil, R. Erban, M. Giles, P.K. Maini, STOCHSIMGPU: parallel stochastic simulation for the systems biology toolbox 2 for MATLAB. Bioinformatics **27**(8), 1170–1171, 04 (2011)
[7.48] D.D. Agostino, G. Pasquale, A. Clematis, C. Maj, E. Mosca, L. Milanesi, I. Merelli, Parallel solutions for voxel-based simulations of reaction-diffusion systems. BioMed Res. Int. **2014**, 10 (2014)
[7.49] L. Dematté, T. Mazza, On parallel stochastic simulation of diffusive systems, in *Computational Methods in Systems Biology* ed. by M. Heiner, A.M. Uhrmacher (Springer, Berlin, Heidelberg, 2008), pp. 191–210
[7.50] M. Creutz, Microcanonical Monte Carlo simulation. Phys. Rev. Lett. **50**(19), 1411–1414, 05 (1983)
[7.51] F. Chen, L. Lovász, I. Pak, Lifting Markov chains to speed up mixing (1999)
[7.52] P. Diaconis, S. Holmes, R.M. Neal, Analysis of a nonreversible Markov chain sampler (2000), pp. 726–752
[7.53] T.P. Hayes, A. Sinclair, Liftings of tree-structured Markov chains, in *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*. Lecture Notes in Computer Science, vol. 6302 (Springer, Berlin, Heidelberg, 2010)
[7.54] K.S. Turitsyn, M. Chertkov, M. Vucelja, Irreversible Monte Carlo algorithms for efficient sampling. Phys. D Nonlinear Phenom. **240**(4–5), 410–414, 2 (2011)
[7.55] M. Vucelja, Lifting—a nonreversible Markov chain Monte Carlo algorithm. Am. J. Phys. **84**(12), 958–968 (2016)
[7.56] S. Duane, A.D. Kennedy, B.J. Pendleton, D. Roweth, Hybrid Monte Carlo. Phys. Lett. B **195**(2), 216–222 (1987)
[7.57] A. Bouchard-Côté, S.J. Vollmer, A. Doucet, The bouncy particle sampler: a nonreversible rejection-free Markov chain Monte Carlo method. J. Am. Stat. Assoc. **113**(522), 855–867, 04 (2018)
[7.58] M. Michel, S. Sénécal, *Forward Event-Chain Monte Carlo: a general rejection-free and irreversible Markov chain simulation method*, 02 (2017)
[7.59] J. Bierkens, P. Fearnhead, G. Roberts, *The Zig-Zag Process and Super-Efficient Sampling for Bayesian Analysis of Big Data*, 07 (2016)
[7.60] J. Bierkens, G. Roberts, A piecewise deterministic scaling limit of lifted metropolis-hastings in the curie-weiss model. Ann. Appl. Probab. **27**(2), 846–882 (2017)
[7.61] E.A.J.F. Peters, G. de With, Rejection-free Monte Carlo sampling for general potentials. Phys. Rev. E **85**(2), 026703, 02 (2012)

[7.62] E.P. Bernard, W. Krauth, D.B. Wilson, Event-chain Monte Carlo algorithms for hard-sphere systems. Phys. Rev. E **80**(5), 056704, 11 (2009)
[7.63] E.P. Bernard, W. Krauth, Two-step melting in two dimensions: first-order liquid-hexatic transition. Phys. Rev. Lett. **107**(15), 155704, 10 (2011)
[7.64] S.C. Kapfer, W. Krauth, Sampling from a polytope and hard-disk Monte Carlo. J. Phys. Conf. Ser. **454**(1), 012031 (2013)
[7.65] S.C. Kapfer, W. Krauth, Two-dimensional melting: from liquid-hexatic coexistence to continuous transitions. Phys. Rev. Lett. **114**(3), 035702, 01 (2015)
[7.66] A. Mendoza-Coto, R. Díaz-Méndez, G. Pupillo, Event-driven Monte Carlo: exact dynamics at all time scales for discrete-variable models. EPL (Europhys. Lett.) **114**(5), 50003 (2016)
[7.67] M. Michel, J. Mayer, W. Krauth, Event-chain monte carlo for classical continuous spin models. EPL (Europhys. Lett.) **112**(2), 20003 (2015)
[7.68] Y. Nishikawa, M. Michel, W. Krauth, K. Hukushima, Event-chain algorithm for the Heisenberg model: evidence for $z \simeq 1$ dynamic scaling. Phys. Rev. E **92**(6), 063306, 12 (2015)
[7.69] M. Michel, S.C. Kapfer, W. Krauth, Generalized event-chain Monte Carlo: constructing rejection-free global-balance algorithms from infinitesimal steps. J. Chem. Phys. **140**(5), 054116 (2014)
[7.70] T.A. Kampmann, H.-H. Boltz, J. Kierfeld, Parallelized event chain algorithm for dense hard sphere and polymer systems. J. Comput. Phys. 0021-9991 **281**(C), 864–875 (2015)

Chapter 8

[8.1] P.G. de Gennes, F. Brochard-Wyart, D. Quéré, *Capilllarity and Wetting Phenomena: Drops, Bubbles, Pearls, Waves* (Springer, New York, 2003)
[8.2] S. Dietrich, in *Phase Transitions and Critical Phenomena XII*, ed. by C. Domb, J.L. Lebowitz (Academic Press, London, 1988), p. 1
[8.3] K. Binder, D.P. Landau, M. Müller, J. Stat. Phys. **110**, 1411 (2003)
[8.4] D. Bonn, J. Eggers, J. Indekeu, J. Meunier, E. Rolley, Rev. Mod. Phys. **81**, 739 (2009)
[8.5] R.B. Potts, Proc. Camb. Philos. Soc. **48**, 106 (1952)
[8.6] F.Y. Wu, Rev. Mod. Phys. **54**, 235 (1982)
[8.7] D.P. Landau, K. Binder, Phys. Rev. B **17**, 2328 (1978)
[8.8] W. Selke, Phys. Rev. E **83**, 042102 (2011)
[8.9] S. Hu, S.-H. Tsai, D.P. Landau, Phys. Rev. E **89**, 032118 (2014)
[8.10] K. Binder, K. Vollmayr, H.-P. Deutsch, J.D. Reger, M. Scheucher, D.P. Landau, Int. J. Mod. Phys. C **3**, 1025 (1992)
[8.11] K. Vollmayr, J.D. Reger, M. Scheucher, K. Binder, Z. Phys. B Condens. Matter **91**, 113 (1993)
[8.12] J. Xu, S.H. Tsai, D.P. Landau, K. Binder, Phys. Rev. E **99**, 023309 (2019)
[8.13] K. Binder, Z. Phys. B **43**, 119 (1981)
[8.14] C.E. Fiore, M.G.E. da Luz, Phys. Rev. Lett. **107**, 230601 (2011)
[8.15] C.E. Fiore, M.G.E. da Luz, J. Chem. Phys. **138**, 014105 (2013)
[8.16] T. Vogel, Y.W. Li, T. Wüst, D.P. Landau, Phys. Rev. E **90**, 023302 (2014)
[8.17] E. Marinari, G. Parisi, Europhys. Lett. **19**, 451 (1992)
[8.18] A. Valentin, C.J. Da Silva, C.E. Fiore, Comput. Phys. Commun. **196**, 213 (2015)
[8.19] O. Philipsen, C. Pinke, Phys. Rev. D **89**, 094504 (2014)
[8.20] C. Czaban, F. Cuteri, O. Philipsen, C. Pinke, A. Sciarra, in *NIC Symposim 2016, Proceedings*, ed. by K. Binder, M. Müller, M. Kremer, A. Schnurpfeil (Research Center Jülich, 2016), p. 31
[8.21] A. Ukawa, J. Stat. Phys. **160**, 1081 (2015)
[8.22] M.E. Fisher, Rev. Mod. Phys. **46**, 597 (1974)

[8.23] V.P. Priman, A. Aharony, P.C. Hohenberg, in *Phase Transitions and Critical Phenomena*, vol. 14, ed. by C. Domb, J.L. Lebowitz (Academic Press, London, 1991), p. 1
[8.24] Y. Imry, S.K. Ma, Phys. Rev. Lett. **35**, 1399 (1975)
[8.25] M. Schwartz, Europhys. Lett. **15**, 777 (1991)
[8.26] T. Nattermann, in *Spin Glasses and Random Fields*, ed. by A.P. Young (World Scientific, Singapore, 1998)
[8.27] K. Eichhorn, K. Binder, Europhys. Lett. **30**, 331 (1995)
[8.28] R.L.C. Vink, K. Binder, H. Löwen, Phys. Rev. Lett. **97**, 230603 (2006)
[8.29] R.L.C. Vink, K. Binder, H. Löwen, J. Phys. Condens. Matter **20**, 404222 (2008)
[8.30] R.L.C. Vink, T. Fischer, K. Binder, Phys. Rev. E **82**, 051134 (2010)
[8.31] K. Binder, Z. Phys. B **50**, 343 (1983)
[8.32] K. Binder, Z. Phys. B **61**, 33 (1985)
[8.33] A. Statt, A. Winkler, P. Virnau, K. Binder, J. Phys. Condens. Matter **24**, 464122 (2012)
[8.34] A. Winkler, A. Statt, P. Virnau, K. Binder, Phys. Rev. E **87**, 032307 (2013)
[8.35] S. Asakura, F. Oosawa, J. Chem. Phys. **22**, 1255 (1954)
[8.36] H.N.W. Lekkerkerker, R. Tuinier, *Colloids and the Depletion Interaction* (Springer, Dordrecht, 2011)
[8.37] K. Binder, P. Virnau, A. Statt, J. Chem. Phys. **141**, 140901 (2014)
[8.38] W. Selke, J. Stat. Phys. **56**, 609 (1989)
[8.39] M.J.P. Nijmeijer, C. Bruin, A.B. van Woerkom, A.F. Bakker, J.M.J. van Leeuwen, J. Chem. Phys. **96**, 565 (1992)
[8.40] A. Milchev, K. Binder, J. Chem. Phys. **114**, 8610 (2001)
[8.41] T. Ingebrigtsen, S. Toxvaerd, J. Phys. Chem. C **111**, 8518 (2007)
[8.42] D. Winter, P. Virnau, K. Binder, Phys. Rev. Lett. **103**, 225703 (2009)
[8.43] D. Winter, P. Virnau, K. Binder, J. Phys. Condens. Matter **21**, 464118 (2009)
[8.44] S.K. Das, K. Binder, EPL **92**, 26006 (2010)
[8.45] S.K. Das, K. Binder, Mol. Phys. **109**, 1043 (2011)
[8.46] K. Binder, B. Block, S.K. Das, P. Virnau, K. Binder, J. Stat. Phys. **144**, 690 (2011)
[8.47] D. Deb, A. Winkler, P. Virnau, K. Binder, J. Chem. Phys. **136**, 134710 (2012)
[8.48] M.L. Trobo, E.V. Albano, K. Binder, Phys. Rev. E **93**, 052805 (2016)
[8.49] J.S. Rowlinson, B. Widom, *Molecular Theory of Capillarity* (Clarendon, Oxford, 1982)
[8.50] J.O. Indekeu, Int. J. Mod. Phys. B **8**, 309 (1994)
[8.51] L. Schimmele, M. Napiorkowski, S. Dietrich, J. Chem. Phys. **127**, 164715 (2007)
[8.52] R.D. Gretz, J. Chem. Phys. **45**, 3160 (1966)
[8.53] G. Navascues, P. Tarazona, J. Chem. Phys. **75**, 2441 (1981)
[8.54] S.K. Singha, P.K. Das, B. Maiti, J. Chem. Phys. **142**, 104706 (2015)
[8.55] P.C. Hohenberg, B.I. Halperin, Rev. Mod. Phys. **49**, 435 (1977)
[8.56] E.V. Albano, K. Binder, Phys. Rev. E **85**, 061601 (2012)
[8.57] D.B. Abraham, Phys. Rev. Lett. **44**, 1165 (1980)
[8.58] D.B. Abraham, E.R. Smith, J. Stat. Phys. **4**3, 621 (1986)
[8.59] D.B. Abraham, in *Phase Transitions and Critical Phenomena*, vol. X, ed. by C. Domb, J.L. Lebowitz (Academic Press, New York, 1986), p. 1
[8.60] E. Brezin, B.I. Halperin, S. Leibler, Phys. Rev. Lett. **50**, 1387 (1983)
[8.61] K. Binder, D.P. Landau, D.M. Kroll, Phys. Rev. Lett. **56**, 2272 (1986)
[8.62] K. Binder, D.P. Landau, Phys. Rev. B **37**, 1745 (1988)
[8.63] K. Binder, D.P. Landau, S. Wansleben, Phys. Rev. B **40**, 6971 (1989)
[8.64] A.O. Parry, C. Rascon, N.R. Bernardino, J.M. Romero-Enrique, Phys. Rev. Lett. **100**, 136105 (2008)
[8.65] L. Pang, D.P. Landau, K. Binder, Phys. Rev. Lett. **106**, 236102 (2011)
[8.66] E.V. Albano, K. Binder, Phys. Rev. Lett. **109**, 036101 (2012)
[8.67] P. Bryk, K. Binder, Phys. Rev. E **88**, 030401 (R) (2013)
[8.68] A.O. Parry, R. Evans, Phys. Rev. Lett. **64**, 439 (1990)
[8.69] A.O. Parry, R. Evans, Phys. A **181**, 250 (1992)
[8.70] K. Binder, D.P. Landau, A.M. Ferrenberg, Phys. Rev. Lett. **74**, 298 (1995)

[8.71] K. Binder, R. Evans, D.P. Landau, A.M. Ferrenberg, Phys. Rev. E **53**, 5023 (1996)
[8.72] M. Blume, Phys. Rev. **141**, 517 (1996)
[8.73] H.W. Capel, Physica **32**, 966 (1966)
[8.74] K. Binder, J.S. Wang, J. Stat. Phys. **55**, 87 (1989)
[8.75] A. Milchev, M. Müller, K. Binder, D.P. Landau, Phys. Rev. Lett. **90**, 136101 (2003)
[8.76] A. Milchev, M. Müller, K. Binder, D.P. Landau, Phys. Rev. E **68**, 031601 (2003)
[8.77] E.V. Albano, A. de Virgiliis, M. Müller, K. Binder, J. Phys. Condens. Matter **15**, 333 (2003)
[8.78] A. Milchev, M. Müller, K. Binder, Phys. Rev. E **72**, 031603 (2005)
[8.79] R.M. Hornreich, M. Luban, S. Shtrikman, Phys. Rev. Lett. **35**, 1678 (1975)
[8.80] W. Selke, Phys. Rep. **70**, 213 (1988)
[8.81] B. Schmittmann, R.K.P. Zia, in *Phase Transitions and Critical Phenomena*, vol. 17, ed. by C. Domb, J.L. Lebowitz (Academic Press, London, 1995), p. 1
[8.82] S. Katz, J.L. Lebowitz, H. Spohn, Phys. Rev. B **28**, 1655 (1984)
[8.83] K.T. Leung, Phys. Rev. Lett. **66**, 453 (1991)
[8.84] S. Caracciolo, A. Gambassi, M. Gubinelli, A. Pelissetto, J. Stat. Phys. **115**, 281 (2004)
[8.85] A.C. Zettlemoyer (ed.), *Nucleation* (Marcel Dekker, New York, 1969)
[8.86] D. Kashchiev, *Nucleation, Basic Theory with Applications* (Butterworth-Heinemann, Oxford, 2000)
[8.87] K.F. Kelton, A.I. Greer, *Nucleation* (Pergamon, Oxford, 2009)
[8.88] K. Binder, P. Virnau, J. Chem. Phys. **145**, 211701 (2016)
[8.89] F. Schmitz, P. Virnau, K. Binder, Phys. Rev. E **90**, 012128 (2014)
[8.90] S.K. Das, J. Horbach, K. Binder, M.E. Fisher, J.V. Sengers, J. Chem. Phys. **125**, 024506 (2006)
[8.91] A. Winkler, D. Wilms, P. Virnau, K. Binder, J. Chem. Phys. **133**, 164702 (2010)
[8.92] K. Binder, Z. Phys. B **45**, 61 (1981)
[8.93] E. Bürkner, D. Stauffer, Z. Phys. B **53**, 241 (1983)
[8.94] M. Hasenbusch, K. Pinn, Phys. A **192**, 342 (1993)
[8.95] M. Hasenbusch, K. Pinn, Phys. A **203**, 189 (1993)
[8.96] K.K. Mon, S. Wansleben, D.P. Landau, K. Binder, Phys. Rev. B **39**, 7089 (1989)
[8.97] M. Hasenbusch, S. Meyer, M. Pütz, J. Stat. Phys. **85**, 383 (1996)
[8.98] J.D. Weeks, in *Ordering in Strongly Fluctuating Condensed Matter Systems*, ed. by T. Riste (Plenum Press, New York, 1980), p. 222
[8.99] H. van Beijeren, I. Nolden, *Structure and Dynamics of Surfaces II*, ed. by W. Schommers, P. Blankenhagen (Springer, Berlin, 1987), p. 259
[8.100] A.M. Ferrenberg, D.P. Landau, Phys. Rev. B **44**, 5081 (1991)
[8.101] E. Bittner, A. Nussbaumer, W. Janke, Nucl. Phys. B **820**, 694 (2009)
[8.102] B.J. Block, S. Kim, P. Virnau, K. Binder, Phys. Rev. E **90**, 062106 (2014)
[8.103] M.P.A. Fisher, D.S. Fisher, J.D. Weeks, Phys. Rev. Lett. **48**, 368 (1982)
[8.104] V. Priman, Int. J. Mod. Phys. C **3**, 857 (1992)
[8.105] K. Binder, Phys. Rev. A **25**, 1699 (1982)
[8.106] L. Onsager, Phys. Rev. **65**, 117 (1944)
[8.107] R.K.P. Zia, J.E. Avron, Phys. Rev. B **25**, 2042 (1982)
[8.108] C. Rottman, M. Wortis, Phys. Rep. **103**, 59 (1984)
[8.109] F. Schmitz, P. Virnau, K. Binder, Phys. Rev. E **87**, 053302 (2013)
[8.110] W.K. Burton, N. Cabrera, F.C. Frank, Trans. R. Soc. **243**, 299 (1951)
[8.111] H.J. Leamy, G.H. Gilmer, J. Cryst. Growth **24/25**, 499 (1974)
[8.112] Y. Saito, H. Müller-Krumbhaar, in *Applications of the Monte Carlo Method in Statistical Physics*, ed. by K. Binder (Springer, Berlin, 1984) (Chapter 7)
[8.113] K.K. Mon, D.P. Landau, D. Stauffer, Phys. Rev. B **42**, 545 (1990)
[8.114] F. Schmid, K. Binder, Phys. Rev. B **46**, 13565 (1992)
[8.115] A. Hashibon, J. Adler, G. Baum, S.G. Lipson, Phys. Rev. B **58**, 4120 (1998)
[8.116] J.D. Weeks, J. Chem. Phys. **67**, 3106 (1977)
[8.117] R.L.C. Vink, J. Horbach, K. Binder, J. Chem. Phys. **122**, 134905 (2005)
[8.118] A. Werner, F. Schmid, M. Müller, K. Binder, Phys. Rev. E **59**, 728 (1999)

[8.119] J.J. Hoyt, M. Asta, A. Karma, Phys. Rev. Lett. **86**, 5530 (2001)
[8.120] A. Asta, J.J. Hoyt, A. Karma, Phys. Rev. B **66**, 100101 (2002)
[8.121] J.R. Morris, Phys. Rev. B **66**, 144104 (2002)
[8.122] T. Zykova-Timan, R.E. Rozas, J. Horbach, K. Binder, J. Phys. Condens. Matter **21**, 064102 (2009)
[8.123] T. Zykova-Timan, J. Horbach, K. Binder, J. Chem. Phys. **133**, 014705 (2010)
[8.124] R.E. Rozas, J. Horbach, EPL **93**, 2006 (2011)
[8.125] D. Jasnow, Rep. Prog. Phys. **47**, 1059 (1984)
[8.126] F. Schmid, K. Binder, Phys. Rev. B **46**, 13553 (1992)
[8.127] F. Höfling, S. Dietrich, Europhys. Lett. **109**, 46002 (2015)
[8.128] A.O. Parry, C. Rascon, R. Evans, J. Phys. Condens. Matter **28**, 244013 (2016)
[8.129] E. Chacon, P. Tarazona, J. Phys. Condens. Matter **28**, 244014 (2016)
[8.130] I. Potoff, A.Z. Panagiotopoulos, J. Chem. Phys. **112**, 6411 (2000)
[8.131] B.M. Mognetti, P. Virnau, L. Yelash, W. Paul, K. Binder, M. Müller, L.G. MacDowell, J. Chem. Phys. **130**, 044101 (2009)
[8.132] R.L.C. Vink, J. Horbach, J. Phys. Condens. Matter **16**, 3807 (2004)
[8.133] R.C.L. Vink, T. Schilling, Phys. Rev. E **71**, 011401 (2005)
[8.134] M.P. Gelfand, M.E. Fisher, Phys. A **166**, 1 (1990)
[8.135] F. Schmitz, P. Virnau, J. Chem. Phys. **142**, 144108 (2015)
[8.136] P. Virnau, M. Müller, J. Chem. Phys. **120**, 10925 (2004)
[8.137] P. Virnau, F. Schmitz, K. Binder, Mol. Simul. **42**, 549 (2016)
[8.138] M.J.P. Nijmeijer, A.F. Bakker, C. Bruin, J.H. Sikkenk, J. Chem. Phys. **89**, 3789 (1988)
[8.139] I.J. Chen, J. Chem. Phys. **103**, 10214 (1995)
[8.140] D. Deb, D. Wilms, A. Winkler, P. Virnau, K. Binder, Int. J. Mod. Phys. C **23**, 1240011 (2012)
[8.141] C. Chatelain, J. Stat. Mech. **0704**, P04011 (2007)
[8.142] M. Caselle, G. Costagliola, A. Nada, M. Panero, A. Toniato, arXiv:1604.05544v1
[8.143] C. Jarzynski, Phys. Rev. Lett. **78**, 2690 (1997)
[8.144] R.L. Davidchack, J. Chem. Phys. **133**, 234701 (2010)
[8.145] J.R. Espinosa, C. Vega, E. Sanz, J. Chem. Phys. **141**, 134709 (2014)
[8.146] J.E. Avron, J.E. Taylor, R.K.P. Zia, J. Stat. Phys. **33**, 493 (1983)
[8.147] M. Heni, H. Löwen, Phys. Rev. E **60**, 7057 (1999)
[8.148] B.B. Laird, R.L. Davidchack, J. Chem. Phys. **132**, 204101 (2010)
[8.149] A. Fortini, M. Dijkstra, M. Schmidt, P.P.F. Wessels, Phys. Rev. E **71**, 051403 (2005)
[8.150] E.M. Grzelak, J.R. Errington, J. Chem. Phys. **132**, 224702 (2010)
[8.151] J.R. Errington, Phys. Rev. E **67**, 012102 (2003)
[8.152] G.A. Voth (ed.), *Coarse-Graining of Condensed Phase and Biomolecular Systems* (CRC Press, Boca Raton, 2009)

Appendix

[A.1] J. Hoshen, R. Kopelman, Phys. Rev. B **14**, 3438 (1976)

Index

K. Binder and D. W. Heermann, *Monte Carlo Simulation in Statistical Physics*,
Graduate Texts in Physics, https://doi.org/10.1007/978-3-030-10758-1